

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2004, 19:37:33 ; Search time 2187 Seconds

(without alignments)  
11831.638 Million cell updates/sec

Title: US-09-613-486-14

Perfect score: 597  
Sequence: 1 atgagatgcagtcgcgcacacg.....gtagaagggaattatgttaa 597

Scoring table: IDENTITY\_NUC  
Gapop 10.0 ; Gapext 1.0

Searched: 3470272 seqs, 2167516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_da:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_pl:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_un:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_inu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pln:\*  
35: em\_hcg\_rod:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vrt:\*  
38: em\_sy:\*  
39: em\_hlgo\_hum:\*  
40: em\_hlgo\_mus:\*  
41: em\_hlgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	597	100.0	597	6 AR138306	AR138306 Sequence
2	597	100.0	15000	6 AF039204	AF039204 Grapevine
3	597	100.0	15500	6 AR138299	AR138299 Sequence
4	593.8	99.5	8590	14 GLA4131	Y14131 Grapevine 1
5	374.6	62.7	16527	14 AF314061	AF314061 Grapevine
6	86.6	14.5	10545	14 BY151931	US1931 Beet yellow
7	70.4	11.8	15468	14 AF190581	AF190581 Beet yell
8	65.6	11.0	15468	14 AF056575	AF056575 Beet yell
9	64.6	10.8	2724	14 BYVCPG	M59452 Beet yellow
10	64	10.7	612	14 BYU71295	U71295 Beet yellow
11	63.8	10.7	6746	14 CLEV3PH	X53462 Beet yellow
12	63.8	10.7	15480	14 BYVUA	X73476 Beet yellow
13	63	10.6	7823	6 A41914	A41914 Sequence 1
14	63	10.6	5960	14 BYVMBPA	X73475 Beet yellow
15	37.8	6.3	2000	6 AX655393	AX655393 Sequence
16	37.8	6.3	151761	9 HSJ493D19	AL096868 Human DNA
17	37	6.2	145710	8 AP004765	AP004765 Oryza sat
18	36.8	6.2	1335	6 AX607511	AX607511 Sequence
19	36.8	6.2	29435	1 AE014256	AE014256 Streptococ
20	36.8	6.2	34838	6 AX602189	AX602189 Sequence
21	36.8	6.2	113050	1 SAG76851	AL76851 Streptococ
22	36.8	6.2	145254	2 AC114681	AC114681 Homo sapi
23	36.8	6.2	164155	9 AC000385	AC000385 Homo sapi
24	36.8	6.2	174171	1 AC068587	AC068587 Homo sapi
25	36.6	6.1	19103	1 AE007030	AE007030 Mycobacte
26	36.6	6.1	218193	2 AC145974	AC145974 Gallus ga
27	36.6	6.1	300050	1 BX248329	BX248329 Mycobacte
28	36.6	6.1	347456	1 BX842577	BX842577 Mycobacte
29	36.4	6.1	87942	9 HS31006	AL035593 Human DNA
30	36.4	6.1	168654	9 AC064878	AC064878 Homo sapi
31	36.2	6.1	625	8 AF462438	AF462438 Scleropha
32	35.8	6.0	142280	2 AC102011	AC102011 Mus muscu
33	35.8	6.0	198882	2 AC110577	AC110577 Rattus no
34	35.8	6.0	206057	2 AC106183	AC106183 Rattus no
35	35.8	6.0	211062	2 AC126868	AC126868 Rattus no
36	35.6	6.0	237979	2 AC121719	AC121719 Rattus no
37	35.6	6.0	593	6 AR237648	AR237648 Sequence
38	35.6	6.0	155780	9 HS336012	AL513498 Homo sapi
39	35.6	6.0	175931	9 HS736	AL163193 Homo sapi
40	35.4	5.9	179342	9 BX255923	BX255923 Human DNA
41	35.2	5.9	224166	2 AC110955	AC110955 Rattus no
42	35.2	5.9	264547	2 AC094432	AC094432 Rattus no
43	35	5.9	123210	2 AC136039	AC136039 Rattus no
44	35	5.9	141773	2 AC069348	AC069348 Homo sapi
45	35	5.9	160930	9 AC010093	AC010093 Homo sapi

#### ALIGNMENTS

RESULT 1  
AR138306  
LOCUS AR138306 597 bp DNA  
DEFINITION Sequence 14 from patent US 6197948.  
ACCESSION AR138306  
VERSION AR138306.1 GI:14479815  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 597)  
AUTHORS Zhu,H.-Y., Ling,K.-S. and Gonsalves,D.  
TITLE Grapevine leafroll virus (type 2) proteins and their uses  
JOURNAL Patent: US 6197948-A 14 06-MAR-2001;  
FEATURES Location/Qualifiers

source	1..597	/organism="unknown"
ORIGIN	/mol_type="unassigned DNA"	
Query Match	100.0%; Score 597; DB 6; Length 597;	
Best Local Similarity	100.0%; Pred. No. 4,7e-182;	
Matches	597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATGAGTGTGATGTCGACAGCAACCTTGAACCTGCTGATTAACCAACGCGCTTACTCTA 60	
DB	1 ATGAGTGTGATGTCGACAGCAACCTTGAACCTGCTGATTAACCAACGCGCTTACTCTA 60	
QY	61 AATGTCGTGACACAGAACCTTTTATCTGCTGAAGTTGAAAAAATGTTGTGCAAAAGG 120	
DB	61 AATGTCGTGACACAGAACCTTTTATCTGCTGAAGTTGAAAAAATGTTGTGCAAAAGG 120	
QY	121 GCTCCTTAACGAGGATATAGAGTGTGTGGTCTACTCTCTTACGCACTCGCGGCAAGA 180	
DB	121 GCTCCTTAACGAGGATATAGAGTGTGTGGTCTACTCTCTTACGCACTCGCGGCAAGA 180	
QY	181 ACCAGTCTCTTAAGTTCAGCGCGCAATTGACAGCTTATATTTTCAAAATGTTTCGA 240	
DB	181 ACCAGTCTCTTAAGTTCAGCGCGCAATTGACAGCTTATATTTTCAAAATGTTTCGA 240	
QY	241 GAGAGCAATGTGTAGTATACAGAGGTGACCTTAAAGAGTATCGACGGGTGTGCGCT 300	
DB	241 GAGAGCAATGTGTAGTATACAGAGGTGACCTTAAAGAGTATCTCAACGGGTGTGCGCT 300	
QY	301 CTCAGTACGTTCACTAATAAATTGAAAGCTTGTGCTGATCTTCACTGAGGCTTACGTT 360	
DB	301 CTCAGTACGTTCACTAATAAATTGAAAGCTTGTGCTGATCTTCACTGAGGCTTACGTT 360	
QY	361 GACTTTGTATCGGCTATTAAGCAAAATTAACCCCACTCAACCGCGGCGGAGTGGGG 420	
DB	361 GACTTTGTATCGGCTATTAAGCAAAATTAACCCCACTCAACCGCGGCGGAGTGGGG 420	
QY	421 ATTCCAGTGAAGATTGCTACTTACGTCGAGATTTTCTGGGATCTTGGCCGAAGCTCTCT 480	
DB	421 ATTCCAGTGAAGATTGCTACTTACGTCGAGATTTTCTGGGATCTTGGCCGAAGCTCTCT 480	
QY	481 GAATTACGCAAAATGAGAGATTTGCGCAGATATGACGCTTAAAACTGAAGTGA 540	
DB	481 GAATTACGCAAAATGAGAGATTTGCGCAGATATGACGCTTAAAACTGAAGTGA 540	
QY	541 GTGCTAAATACACAGTGAACATCTCCGTCAGTGAAGTGAAGGAAATTTGTA 597	
DB	541 GTGCTAAATACACAGTGAACATCTCCGTCAGTGAAGGAAATTTGTA 597	
RESULT 2	15000 bp RNA linear VRL-22-MAY-1998	
AF039204		
LOCUS		
DEFINITION	Grapevine leafroll-associated virus 2 methyltransferase/helicase polyprotein gene, partial cds; and RNA-dependent RNA polymerase, putative transmembrane small hydrophobic protein, 65 kDa chaperone protein, 63 kDa protein, 25 kDa diverged coat protein, 22 kDa coat protein, 19 kDa protein, and 24 kDa protein genes, complete cds.	
ACCESSION	AF039204	
VERSION	AF039204.1 GI:3123909	
KEYWORDS		
SOURCE		
ORANISM	Grapevine leafroll-associated virus 2	
	Grapevine leafroll-associated virus 2	
	Vinifera; ssRNA positive-strand viruses, no DNA stage;	
	Closteroviridae; Closterovirus.	
REFERENCE	1 (bases 1 to 15000)	
AUTHORS	Zhu,H.Y., Ling,K.S., Gosczyński,D.E., McFerson,J.R. and Gonsalves,D.	
TITLE	Nucleotide sequence and genome organization of grapevine leafroll-associated virus-2 are similar to beet yellows virus, the closterovirus type member	
JOURNAL	J. Gen. Virol. 79 (Pt 5), 1289-1298 (1998)	
MEDLINE	98264507	

PUBMED	9603345
REFERENCE	2 (bases 1 to 15000)
AUTHORS	Zhu,H.Y., Ling,K.S. and Gonsalves,D.
JOURNAL	Submitted (18-DEC-1997) Plant Pathology, Cornell University, New York State Agricultural Experiment Station, Geneva, NY 14456, USA
FEATURES	location/Qualifiers
source	1..15000
CDS	/organism="Grapevine leafroll-associated virus 2"
	/mol_type="genomic RNA"
	/db_xref="taxon:64003"
	<1..7423
	/function="replication and unknown functions"
	/note="larger than 277 kDa; ORF1a; contains domains for two papain-like leader proteases, a methyltransferase and a helicase; identified by sequence comparison"
	/codon_start=2
	/product="methyltransferase/helicase polyprotein"
	/protein_id="AAC40855.1"
	/db_xref="GI:3123910"
	/translation="ADVYAMLEKYCGKFPVLMSRVTPDRCYLAMRYLCAFCYOR PFRSDYALGMPPTVRLAACEKNPGEVACGIALRGYTRSNVGYCDSAYVYK NLSGRIGGSPDTSITLVITVIGSLGCPKINAFAPLDIRVYVSGIOSS VKTREDFARIVVEERAAAGSVOPROKIDEGRCVRVSGSHLVNLEVRKVA AGLIRFVGVDHFRRFSTCGHRLVWRSSRVCLEIYSPKNTFLRYVLCGSG YAMSFRAAGRPVLVMTIRIYRPNGPCVLAACRACAKILRGSDYAPDGAAPTAA KLRNRYFSLGRSLGLNLYGATYRQVPHCDYDAKFIKRLMSAYIAGDSQRETV PSDITPAMQKITEAYIDRLVGGTSLKLSIEKOLIDKNDVSLKDRITVVPFY MSEATQNSLTRYFPOPELKFSSHSHDPAASAILENTVLNCGNSVDIGGCP FLHSHQRTORVAVCPVLDGKQORVVRDIOYSNVRIGEDDKILBGRNIDICHP GACDHESSAMMYQVADALYEICAMIKKXRIYLTMTVPGFLEGEVYVESLD CELEVDADVVMYKFGSSCYEHLKILKIDMTPEYTLGGLFSPSEVMEYRVMYVF KITKSEVPSISCTKLIRYRANSDVYVQIPRPKRMCPLGVDITLYDSKVSIV PDVYVNCSAVNSKPEYMWSPKSKSRVILSGITHKVDNLKLVESPAAMLAS GRRSLASLYLAKNLSHSGDGSFEIATSPYLREKIRMTINPRLIOLYRVA LDVPSLDSTLSITDPAECVATLEEDLQCRABENKIRLUGDSIAKRLASIE VVDIDKSPKQVNSSSBNADKREVQPGIRGSRNGVVEFLHVVDSLRLEKTA TQOQRIKSYRFLDSAVSFLDYNNDLSEILRVISEGSCWFAFLARGDSVRSRA VCAVKEVATSCANASVAKAWITPAAVCAWMTNSCGFSGDREYKSYLHRYQVLF DITFEDSSSYLDEIVLSAICGALVTLPSSGSSISLANFLQITKGFLEVVANVVR VTHGLSTATDGVIRFSQAIASHLVNTGNVAYOSAFIAGVPLVYKCVSLIFIL RLHYTSGFIFKHGISSEFSLIKLEKLVDELKSIIGVFPDKNKVFPEAQEAR TVMQVPAVAVDALKSAAGKILYNTSPSTSPKDGSGSDGACVSGCOBGRGPK GGSGYFSLTAFSRIIMGARRLLYKVAHEFCGLTPELSEKHEFTTRPGKRTV GYDPLGTGVTLSIECQIIEBLDDKQDMSALTDWNTFERSSEPLASTIEP PAGLRGSGKNTAILALEYAHNFRIVAYASCKSPFLPFAVLELSALIEKFEVER KQQLVAIAEYVQFLRSMRALGLNNEFVYKSPADLLPALMRKVGSTLASVYRPL RGSFYVCAEERKEFPALVLCIGSLPFAVIRVGAACAEELVSAPRERIKIETPL OKYVLSINFPCHPSSDDVDSASAGLKGASNTMLPHILVRLASALLSGWELKLL LSHNHLLFLCFALVDVNVLLIKVGLGSLFFQPIESLFAMLLQDPDFVSEGLVA FSPFLKSPRAPALIKGFPECVANSTYKTRILRCFVQVKKLRKRGARAGRGH RQKAVDVIDSNVVDGVRLSVYMGVGEALRTETRLITEDLSAVIKLARPRRND DDEFTPRAEQMOQVTTTCSRANAGLLEAVIYEDALIKESKTYNEMVYKGGSTT SEBAVALSDDEAVELISVADERDSDPTVATIRYNFLKNSSEFPQPIVDDKDTG GLNVAAREYVQELALFEIHSKICTYDQIRIYNFVRSAVCESDQLVNRKNGST VQGEKVRALIKDFHDFLPDKLISINRRKGVNLDHNLAFASMLPAGVFPSS EYFTNVSVDILILEAPPGGKTTLLISGLFVFKGVSITWILNANSQVEILKYVE KEVNSIECQRDKRSPKKSIVITIDYAMHRGDAVLFIDEFMHAGSVAICIFP TGHKVMIEGDSROIHYIERNELDKICGLDRLVDPOCRVYGISVRCMGVACMS TYVKNLITVYGESRSGKSMRINRINSNVDPFVPSGTPICOMQSEKTLISKPIRG LTKVNLITVYHESQETTYARNVLVRLKQEDBEPFSGIHTITATLSRHTDSITVYLA RQDATCDAIQKRAELVNRVYFPTSGSVNLNVRKQVDEBNSCKRSLAPLVINDF LNEVPGTAVIDFGLSDAFSTGPCEAGSIVVRNDISSNTIDHKQRY"
CDS	/function="replication"
	/note="RdRp; 52 kDa; similar to RNA polymerases of other closteroviruses; presumably expressed via +1 ribosomal frameshift"
	/codon_start=1
	/product="RNA-dependent RNA polymerase"
	/protein_id="AAC40856.1"
	/db_xref="GI:3123911"
	/translation="SVNSQAIIPRRKPSLOENTLSPFARYNVSFTCDRNTASMGFA

MANCLRCFEDLDFASSLRDVYISITSGIEOMLEKRTPSQIKALMKOVESPLEIDDE  
ICREFKLVKADAKVLDSCLTGHSAQNIIMFKRISINAIFSPINFEVNRIMCLKP  
NIKEFTMTNRDPAVSNIGDDVYHIGEVDSKYDQSGADPAVEVYKELGV  
DEELAIIMMGGERLISANTLDGQFTIENOKRSGASNTIGNSLVTLGILSYDVR  
NBEALVYISGDDSLIFSRSEISYADICTDMGERTKMSPSVYSCRYVMCGKHTP  
FPDPFKLTVKCAVAKEDVSMPLFTFTSPKDLISDRFDERLIQKLEVALKTEVQ  
TGATTLATSVIHCLRSNPLFSKLYPRVGMVFTYSVKALKKSGCSPDSFMPFG  
QAMVWDEE  
8865. .9035  
/note="6 kDa; probably membrane-associated; similar to  
small hydrophobic proteins of other closteroviruses"  
/codon\_start=1  
/product=putative transmembrane small hydrophobic  
protein  
/protein\_id="AAC40857.1"  
/db\_xref="GI:3123912"  
/translation="MNQVLFCECLFLNLDAVAVTFPIFLVFRVVKSPRQGHKAVP  
PVYRGGGFSIVV"  
9051. .10850  
/note="p65; HSP70; similar to heat shock 70 proteins;  
identified by sequence comparison"  
/codon\_start=1  
/product="65 kDa chaperone protein"  
/protein\_id="AAC40858.1"  
/db\_xref="GI:3123913"  
/translation="MVFRLDPTFTSTVYCYKQGVSPFKONNSAYITPYLYLSDS  
NHTFGTEASLMSNLKVGSYFRDLKRWYCGDSSNLDLRLKPHISRLVKISG  
LNEVTSIGNFGTVKSEAHPLGLIALFIAVVISGACGACTGTGTCVPPANYSVQ  
RNFTDQCVLSGYQCYVINEPSPALISACNIGKSNAMAYDFGGGTFIISYR  
NNTFVVRASGDLNLGSDVDRAFLTHLFTSLTSLDNLISNKKESKIDAIIV  
TLTGVRKEDVAVNNKILTSVMLPYVNRILKLESTLSYKMMESRKAIDVLI  
GSSSYLGLADVLTKHQSDRLIRVSDPPAAVAGALYSCGSGGGLIDCAHT  
VALADSSCHQIICAPGAPRPSGSMPIIARVNRKSGQEVFBEZYKCRKNNKIC  
GAKRFIDIGVDSYAPVTFYMDSTISSGAVSFYVRPBGQVSLTPEANSSV  
AAGSRVRELHILNKKVPLGLLHKKARILLFTDEAIRVADSIDADIVKEYKSY  
AASALPDEDEVELLNGKSVQKVLGRSRLIEIPL"  
10777. .12432  
/note="p63; putative heat shock protein 90 homolog"  
/codon\_start=1  
/product="63 kDa protein"  
/protein\_id="AAC40859.1"  
/db\_xref="GI:3123914"  
/translation="MNSYWSLFFKKYGEADMKYLSRSHASHSEIKTLDIRLYG  
GRVKSSEFSAALPNSFEQBLGLFLLSEEVGWSKLCITVEEAYDLINPAAFYTA  
ETSPVYKSGCKYSMEDVNMFMRLISNLNDMLTEQCMSSGSELINDDGFR  
VALTFKDRDPAADTGAANVECRGYDLYVAMSLFEORTKQSGNSLVEKCEYRT  
YLGSTPLPADPRLPLTGILYDRCKEYVNPYSSYKRVNDNRPRLANMPLISDV  
VFQWVAPADVRLLELSSAELTLEVTSLISDSOYVGHILRYESTISDALDALE  
DKKALIKSNPRLSTAQLVNGFCYIGERTQOSVVRPGVYKTPDSVGSFELNMK  
DYKFFDKLQRELPNVSLRQFNGARHAFKIPKNGTISFPIISHLNVPREFWLT  
DYERANRSGLELEILLINISVDVRKICAEBACTNLLSARFSPKSNHNSNQSROE  
RRIKDLVVLKDTLEFQHRKAGWGSRSRDLGSRADHAKGSG"  
12344. .13015  
/note="p25; Cpd; coat protein duplicate"  
/codon\_start=1  
/product="25 kDa diverged coat protein"  
/protein\_id="AAC40860.1"  
/db\_xref="GI:3123915"  
/translation="MNSNTSVPVGGLALETSGVLTTRKEAVDKFFNLKNENYSSV  
DSRLSDSEYKLEKSKESFSELASTTBHFVYHIFLPLICLXSTIEKRYGSH  
TYVVDKTYTVDLAWYPMNMKSLTKRYKCVNGRACCCEDLYLTVAISEMREKTY  
AVMKCLPVKKEYLGADFLSGTSKLSMDHRANYSIYAACNAVDRSAFTGGEKITYSLY  
DLGRY"  
13084. .13680  
/note="p22"  
/codon\_start=1  
/product="22 kDa coat protein"  
/protein\_id="AAC40861.1"  
/db\_xref="GI:3123916"  
/translation="MEIMSDSNLSNLTVDASLNGVDKLLSAEYKMLVQKGAANE  
GIEVFGLLLYALAARTSPKQVADSDVIFNSFGBERVVATGDLKLVLDGCAFLT  
RFTNKLRTFGRTTEAVVDFCIAYKHKLPQLNAAELGIPADBSYLAADFLCTPCLTS

ELQSRKQFASMTALKTGGVAVNPVSNLRQLGRREV"  
13680. .14165  
/note="p19"  
/codon\_start=1  
/product="19 kDa protein"  
/protein\_id="AAC40862.1"  
/db\_xref="GI:3123917"  
/translation="MEDYEEKSESLILRLTNLTMLLVKSDASVELPKLLCGYLRV  
SGRGEVTCNRBELTRDFEGNHHVIRSLIQYDSEAFEEFNNSDGVFPLETGSV  
FMPFLRSETGRBAVRHLRTFPEANPFPGSHCGMEYCLKQVLTETESLIDSCERN  
R"  
14167. .14784  
/note="p24"  
/codon\_start=1  
/product="24 kDa protein"  
/protein\_id="AAC40863.1"  
/db\_xref="GI:3123918"  
/translation="MRVIVSPYEADILKRSITMLRVIDSGLSTKECIKAFSTTRD  
LHCKASYQNGVDGILQORCAERKRLIDTIESNIRLAQPLVREKVAHPRCKDREKY  
AFIRKTYELTGVRERARERKMSLTKTVLNKRSLEMAFMRKAMLELKF  
PVKIFRDLIDVETINELCAEDDVHDKVNEDEHDLQDEC"  
14785. .15000  
/note="similar to 3' UTRs of other monopartite  
closteroviruses"  
3'UTR  
CDS  
ORIGIN  
Query Match 100.0%; Score 597; DB 14; Length 15000;  
Best Local Similarity 100.0%; Pred. No. 7,7e-182;  
Matches 597; Conservative 0; Mismatches 0; Gaps 0;  
1 ATGGAGTGAAGTCCGACAGACGCTTACGACCTGGGATACCGACGCTCTAGTCTA 60  
13084 ATGAGATTGATGTCGACAGACGCTTACGACCTGGGATACCGACGCTCTAGTCTA 13143  
61 AATGGTGTGCAAGAAAGCTTTTATCTGCTGAAGTTGAAAAATGTTGGTACAGAAAGG 120  
13144 AATGGTGTGCAAGAAAGCTTTTATCTGCTGAAGTTGAAAAATGTTGGTACAGAAAGG 13203  
121 GGTCTTACGAGGGGATATGAAGTGTGTGCTTACCTCTTACGACCTGGGACAGAGA 180  
13204 GGTCTTACGAGGGGATATGAAGTGTGTGCTTACCTCTTACGACCTGGGACAGAGA 13263  
181 ACCAGCTCTCTAAGTTCAGCGGCGAAGTTACAGAGTTATATTTCAATAGTTGGGA 240  
13264 ACCAGCTCTCTAAGTTCAGCGGCGAAGTTACAGAGTTATATTTCAATAGTTGGGA 13323  
241 GAGAGGAATGTGTATGATACAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 300  
13324 GAGAGGAATGTGTATGATACAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 13383  
301 CTCACTAGGTTCACTAATTAACCTTGAAGACGTTGGTGTGATCTTCACTGAGAGCTTACGT 360  
13384 CTCACTAGGTTCACTAATTAACCTTGAAGACGTTGGTGTGATCTTCACTGAGAGCTTACGT 13443  
361 GACTTTTATCGCGGTATATAGCACAATTAACCCCAATCAACGCGGCGGGAATGGGG 420  
13444 GACTTTTATCGCGGTATATAGCACAATTAACCCCAATCAACGCGGCGGGAATGGGG 13503  
421 ATTCCAGCTGAAGATTGCTACTTACTGACGAGATTTTGTGGTACTTCCCGGAGCTCTCT 480  
13504 ATTCCAGCTGAAGATTGCTACTTACTGACGAGATTTTGTGGTACTTCCCGGAGCTCTCT 13563  
481 GAATTACGCAAGATGAGAGATGTTGGCGGAGTATGTAAGCTCTTAAAAAATGAAAGGTGGA 540  
13564 GAATTACGCAAGATGAGAGATGTTGGCGGAGTATGTAAGCTCTTAAAAAATGAAAGGTGGA 13623  
541 GTGTAAATACCAAGTGAAGATCTGGCTCAGCTAGGTAGAGGAAGTATATGTA 597  
13624 GTGTAAATACCAAGTGAAGATCTGGCTCAGCTAGGTAGAGGAAGTATATGTA 13680

RESULT 3  
AR138299

LOCUS AR138299 15500 bp DNA linear PAT 16-JUN-2001  
 DEFINITION Sequence 1 from patent US 6197948.  
 ACCESSION AR138299  
 VERSION AR138299.1 GI:14479808  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 15500)  
 AUTHORS Zhu H.-Y., Ling K.-S. and Gonsalves D.  
 TITLE Grapevine leafroll virus (type 2) proteins and their uses  
 JOURNAL Patent: US 6197948-A 1 06-MAR-2001;  
 FEATURES  
 source 1. 15500  
 /organism="Unknown"  
 /mol\_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 597; DB 6; Length 15500;  
 Best Local Similarity 100.0%; Pred. No. 7,7e-182;  
 Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGTTGATGTCGACAGCACTTACGACCTGGTGAATTAACGACGCTTACTCTA 60  
 DB 13584 ATGGAGTTGATGTCGACAGCACTTACGACCTGGTGAATTAACGACGCTTACTCTA 13643

QY 61 AATGGTGTGCAGANAGCTTTTATCTGCTGAAGTGAATAAATGTTGGTCAGAAAGG 120  
 DB 13644 AATGGTGTGCAGANAGCTTTTATCTGCTGAAGTGAATAAATGTTGGTCAGAAAGG 13703

QY 121 GCTCTTAACGAGGGTATGAAGTGTGTGCTTACTCTTACGACCTCGCGGCAAGA 180  
 DB 13704 GCTCTTAACGAGGGTATGAAGTGTGTGCTTACTCTTACGACCTCGCGGCAAGA 13763

QY 181 ACCACGTCCTCTTAAGGTTACGCGGAGATTACAGCTTATATTTCAATTAAGTTTGG 240  
 DB 13764 ACCACGTCCTCTTAAGGTTACGCGGAGATTACAGCTTATATTTCAATTAAGTTTGG 13823

QY 241 GAGGAGATGTGTGAATGAAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 300  
 DB 13824 GAGGAGATGTGTGAATGAAGAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 13883

QY 301 CTCACCTAGTTCACTAATAAAGCTTGAAGCTTGGTGTACTTCTCACTGAGGCTTACGT 360  
 DB 13884 CTCACCTAGTTCACTAATAAAGCTTGAAGCTTGGTGTACTTCTCACTGAGGCTTACGT 13943

QY 361 GACTTTGTATCGGTTAAGCAAAATTAACCCCACTCAACGCGCGGCGGAATTGGGG 420  
 DB 13944 GACTTTGTATCGGTTAAGCAAAATTAACCCCACTCAACGCGCGGCGGAATTGGGG 14003

QY 421 ATTCACAGTGAAGTGTGACTTACTGACAGATTTTCTGGGTACTTCCCGAAGCTCTCT 480  
 DB 14004 ATTCACAGTGAAGTGTGACTTACTGACAGATTTTCTGGGTACTTCCCGAAGCTCTCT 14063

QY 481 GAATTAACGCAAGTGAAGAAATGTGGCGAGTGTACGCTCTAATAAAGTGAAGGGA 540  
 DB 14064 GAATTAACGCAAGTGAAGAAATGTGGCGAGTGTACGCTCTAATAAAGTGAAGGGA 14123

QY 541 TTGGTAATACACAGAGTGAAGTGTGCTGCTAGCTAGTGAAGGAGTGAATGTAA 597  
 DB 14124 TTGGTAATACACAGAGTGAAGTGTGCTGCTAGCTAGTGAAGGAGTGAATGTAA 14180

RESULT 4  
 GLAV131 8590 bp RNA linear VR1\_08-JAN-1998  
 LOCUS Grapevine leafroll-associated virus 2 genes encoding RNA polymerase  
 DEFINITION and coat protein, hsp70, hsp90 gene and ORF2, ORF7 and ORF6.  
 ACCESSION Y14131  
 VERSION Y14131.1 GI:2765391  
 KEYWORDS coat protein, heat shock protein 70-like; ORF1; ORF2; ORF3; ORF4;  
 ORF5; ORF6; ORF7; ORF8; RNA polymerase.  
 SOURCE Grapevine leafroll-associated virus 2

ORGANISM Grapevine leafroll-associated virus 2  
 REFERENCE Viruses; ssRNA positive-strand viruses, no DNA stage;  
 AUTHORS Closteroviridae; Closterovirus.  
 TITLE 1  
 JOURNAL The nucleotide sequence of the 3' terminal region of grapevine  
 2 leafroll associated closterovirus 2  
 Unpublished  
 Abou-Ghanem N.  
 Abou-Ghanem N.  
 Direct Submission  
 Submitted (01-JUL-1997) N. Abou-Ghanem, Universitat degli Studi di  
 Bari, Dipartimento di Protezione delle Pianta e Centro studio  
 virus e virosi delle colture mediterranee, via Amendola 165/A, I  
 70126 Bari, ITALY  
 revised by (3)  
 3 (bases 1 to 8590)  
 REMARK  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 FEATURES  
 source

CDS

1. 8590  
 /organism="Grapevine leafroll-associated virus 2"  
 /mol\_type="genomic RNA"  
 /db\_xref="taxon:64003"  
 /lab\_host="N. benthamiana"  
 1165..2427  
 /note="ORF1"  
 /codon\_start=1  
 /product="RNA polymerase"  
 /protein\_id="CA74561.1"  
 /db\_xref="GI:2369865"  
 /db\_xref="GOA:O39852"  
 /db\_xref="SPTREMBL:O39852"  
 /translation="MFGAMAMNCLRCFQDLDAFSLRNVDYISITRSIGIQLKQKTP  
 SRIRAKMDVESPLEIDELCRFLKMTKVRDKVLDSSCLTKSNAQINMHRKSIYA  
 IESPIENVARIRINCLKPKIKFTFTNDRPAFVNMGGDDIVYIGVDESKYK  
 SODAFVAFSEVMTRELSVDEELAIIMCGSRLEIAATLDQSLFTIENRQKASNT  
 WIGNSLVTLGLSLYVDVNEALYIGDSLLFSRSEISYADICTDGFETKENS  
 SPVYFCKSEFVMOGKHFTFLVLDYKLFVKLGAVKEDVMSDSFETFSFEDLTSFN  
 DERLIQKLAELVALKEYQVTGNTLALSVHICLSNPLSFKLYPRVKMGQVFTSVK  
 KALKSGCSLPSDFMTFPGAVMWVWDE"  
 2491..2661  
 /note="ORF2"  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="CA74562.1"  
 /db\_xref="GI:2369866"  
 /db\_xref="GOA:O39853"  
 /db\_xref="SPTREMBL:O39853"  
 /translation="MNQVLQFECFLNLAFAVTFIFILLVRFVKSFRQKGEAPV  
 PVRGGSSTIV"  
 2677..4476  
 /gene="hsp70"  
 /gene="hsp70"  
 2677..4476  
 /gene="hsp70"  
 /codon\_start=1  
 /product="heat shock protein 70-like"  
 /protein\_id="CA74563.1"  
 /db\_xref="GI:2369867"  
 /db\_xref="GOA:O39854"  
 /db\_xref="SPTREMBL:O39854"  
 /translation="WVFGDLDFGTFSTVVCYKQGRVSEFKONNSAVIPTLYLFSDS  
 NMFTFGYAESLMSNLKVKGSFYRDLKMWGCDSSNLDAYDRKPHYSVFLVKGSG  
 LNEVYIGNFGGVSEAHLPGLTALFTKAVISGAEAFACCTGCTGVCISVANDSVQ  
 NRPFOCVISIGYCVVMINERPSAASACSVKSKSNLANLVDFGGTFFVSIISYR  
 NTFIVRASGSDINLIGSDVDRAPLTHFLFSLTSEPDITLYSNLKESTSTDEIYV  
 TURGVGRKEDVRVKNKILVYMLFTVKTITLESITKSYAKSNBESALYKCDLVLT  
 GSSSYLPGLADVLTRHGSVDRLTRVSDPRAAVAVGCAVYSCLSGSGGLTIDCAAT



VAIADRSCHQIICAPAGADIPSSGSMPLYLAAVNKNNSOREIAVEGEYVCKPKNKIC  
GNIREFDVGWDSYAPVTFYMDVSISSVGVASVVRPBEKQVSLGTAPVANSY  
AAGSRRELIHINNKKVFLGLHHRKARLLFTFDEAIRYADSIADIADLVKEXSY  
AASALPDEDVELLKGKSVOKVLRGSRLEIPL"

gene  
CDS  
/gene=".6058  
/gene="hep90"  
4403..6058  
/codon\_start=1  
/product="heat shock protein 90-like"  
/protein\_id="CA74564.1"  
/db\_xref="GI:2369868"

/db\_xref="SPRMBL:039855"  
/translation="MSNYGMSLFFKPYEADMKYLSIAHSEIKTLPDILY  
GRVKKSEFESALPNSFEOLGLFISEVEGMSKCGITVEAADLTPAPKPTA  
ETGSPVKGQGYQSMEDVNNFRLNSLVNDKMLAQCSLSNSCGELVNDGGRF  
VALTFPDRIADDTGAANVECRGYDLVYAMSLFEORTQSGSNGISLEYKCEYRT  
YLSGDLFPADPRLPLTGLIHIDPCKEYNI FYSYKRVNDPREFLAHPLISDF  
VPCWVAPAPVRLPFLISAELTLEVTSLIDSVVGHILRYVESYSDAIDALE  
DKLEIKSNPRLSTAOJMWGPFYCGEFPNOSRVVORPQVYTPPSVGFEINM  
DVEKFPDKORELPNYSLRQENRGAHAFKIKNGNISFPISRLINVPREFYTLN  
DIFRANRSGLTLEEILINNISVDVRKICARACSTPSARFSPNHSKNQSROE  
RRKIDPLVYLKTLVEFOHRRAGWMSRSTRDLSRADHAKGSG"  
5970..6641  
/note="ORF5  
duplicate gene"  
/codon\_start=1  
/product="coat protein"  
/protein\_id="CA74565.1"  
/db\_xref="GI:2369869"  
/db\_xref="GOA:039856"  
/db\_xref="SPRMBL:039856"  
/translation="MSSNTSVPGLEALETSGLVLTTRKAVDPKNEIKNENYSSV  
DSRLSDSVKVTLEKSKESFSESLASDEHVVHILPELICAKISSEKRYGXT  
TYVDSKRTYVLDAYFNMMKSLTKRYKRVNGIRAFCCCEPLIYITAPINSEKXTK  
AVGMKGLPVGEKYLGLADFLSGTSLKMSDHDRAVSTVAANAVDSAFTEGGERIVSLY  
DLGRY"

CDS

6710..7306  
/note="ORF6"  
/codon\_start=1  
/product="coat protein"  
/protein\_id="CA74566.1"  
/db\_xref="GI:2369870"  
/db\_xref="GOA:039857"  
/db\_xref="SPRMBL:039857"  
/translation="MELMSDSNLNLTIDASSLNGVDKLLSAEYVKMLVOKAPNE  
GIEVFGLLVLAARTSPKVRADSDVFNSSPEERNVVTGDLKKVLDGCAPLT  
RFTNKLRTGRTTEAVYDPCIAVKKKLPOLNAAAEIIGPADSVIADFLGTCKLS  
ELQOSKRMASMAALTBGGVYNTPVSNIRQLGRRWV"  
7306..7791  
/note="ORF7"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CA74567.1"  
/db\_xref="GI:2369871"  
/db\_xref="SPRMBL:039858"  
/translation="MEDYKESBSLILNTNNTMLLVKSDASVELPULICGLIYV  
SGGEVTCNRELTDFEGNHITVRSRIIYDSSAAEEENNSDCCVCFLELGEV  
FMFPLSETKGRAVRLRTFFEANNFPSHCGTWEYCLKQVLTETESLIIDSCERN  
R"

CDS

CDS

7793..8410  
/note="ORF8"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CA74568.1"  
/db\_xref="GI:2369872"  
/db\_xref="SPRMBL:039859"  
/translation="MEVIVSPYAEADILKRTSDMLRNIDSVLSTKCIKAPSTTRD  
LHAKASYOMGVDTGLYORNCAEKRLIDTVESNIRLAQPLVREKMAVHCXQPEELV  
AATRKVELTGVGMEAYKREKRSITKIVLKMSIEMAFVSPAPAKMAEMLKFS  
PVATFDLADVETTLNELCAEDDVHVDKVNEDENHDLQDEC"

ORIGIN

Query Match 99.5%; Score 593.8; DB 14; Length 8590;  
Best Local Similarity 99.7%; Pred. No. 7,7e-161;  
Matches 595; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	ATGAGAGTGAATGTCGACAGCAACCTTTAGCAACCTGGTATTAACCCAGCGCTAGCTA	60
DB	6710	ATGAGAGTGAATGTCGACAGCAACCTTTAGCAACCTGGTATTAACCCAGCGCTAGCTA	6769
QY	61	AATGGTTCGACAGCAAGAGCTTTATCTGCTGAAGTTGAAAAATGTTGGTACAGAAAGG	120
DB	6770	AATGGTTCGACAGCAAGAGCTTTATCTGCTGAAGTTGAAAAATGTTGGTACAGAAAGG	6829
QY	121	GCTCCTAACGAGGGGTATAGAAGTGGTGGTGTACTCTCTTACAGCACTCGCGGACAGA	180
DB	6830	GCTCCTAACGAGGGGTATAGAAGTGGTGGTGTACTCTCTTACAGCACTCGCGGACAGA	6889
QY	181	ACCAAGTCTCCTTAAGGTTACGCGGACAGATTACAGCGTTATATTTCAAAATAGTTTGGGA	240
DB	6890	ACCAAGTCTCCTTAAGGTTACGCGGACAGATTACAGCGTTATATTTCAAAATAGTTTGGGA	6949
QY	241	GAGAGGAATGTGTATGTAACAGAGGGTGACCTTAAGAAAGTACTCGACGGGTGGCCCT	300
DB	6950	GAGAGGAATGTGTATGTAACAGAGGGTGACCTTAAGAAAGTACTCGACGGGTGGCCCT	7009
QY	301	CTCACTAGGTTCACTAATTAACCTTACAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	360
DB	7010	CTCACTAGGTTCACTAATTAACCTTACAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	7069
QY	361	GACCTTGTATGCGGTATAGCAACAATTAACCCCACTCAACCGCGGGGGAATTTGGG	420
DB	7070	GACCTTGTATGCGGTATAGCAACAATTAACCCCACTCAACCGCGGGGGAATTTGGG	7129
QY	421	ATTCAGGTGAAGTGTGTAAGTCTGCTGAGATTTCTGGGTACTTCCCGAAGCTCTCT	480
DB	7130	ATTCAGGTGAAGTGTGTAAGTCTGCTGAGATTTCTGGGTACTTCCCGAAGCTCTCT	7189
QY	481	GAATTAACGACAAAGTAGAAGATGTTCCGAGTAGTAGTACGCTTAAAACTGAAGTGA	540
DB	7190	GAATTAACGACAAAGTAGAAGATGTTCCGAGTAGTAGTACGCTTAAAACTGAAGTGA	7249
QY	541	GTGGTAATTAACAGAGTAGAAGTCTGCTGAGTGTGTAAGTGAAGTGAAGTGAAGTGA	597
DB	7250	GTGGTAATTAACAGAGTAGAAGTCTGCTGAGTGTGTAAGTGAAGTGAAGTGAAGTGA	7306

RESULT 5  
AF314061  
LOCUS  
DEFINITION  
Grapevine rootstock stem lesion associated virus  
polymerase, p6, HSP70-like protein, HSP90-like protein, coat  
protein duplicate, coat protein, p19, and p24 genes, complete cds.  
AF314061  
AF314061.1 GI:24636914

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Grapevine rootstock stem lesion associated virus  
Grapevine rootstock stem lesion associated virus  
Viruses; ssRNA positive-strand viruses, no DNA stage;  
Closteroviridae; Closterovirus.

REFERENCE  
AUTHORS  
TITLE  
Nucleotide sequence of grapevine rootstock stem lesion associated  
virus

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
Direct Submision  
Submitted (16-OCT-2000) Plant Pathology, University of California,  
One Shields Ave., Davis, CA 95616, USA

FEATURES  
source  
1..16527  
/organism="Grapevine rootstock stem lesion associated  
virus"

```

CDS
/mol_type="genomic RNA"
/db_xref="taxon:167634"
107..8902
/note="327 kDa"
/codon_start=1
/product="methyltransferase/helicase polypeptide"
/protein_id="AAN63466.1"
/db_xref="GI:24636918"
/translating="MSLAIISALPSAVAOUSFQOPATYAKSFYMTSLPSLOTCTPQC
EFSPLFICLASFOKIKTFLFSLVRELAFLSELSTGFCCEPIRLGVEHPLFFVFSF
QCGPLSYCLGCVNGIFAVLINISFPLDVALSLNDVDVVTPTVIRSKRAKLNKAF
SIAAARIRSKYKNGEVRISHVQPLAGCLFSONKTELLIIGRVGVARVAT
APVRRGGGGRVPLIPLVAVPFGVRCVAPSCGALACFNKLASVAAANDSAE
KLSAPRRKIOONENSRVAKADVTAASVMAABGVMPRIFFGSGSEPPPISS
SGMARSVLEBPALVTGASAPAKETPRYSYKADPSAPAPREPALPRVAKREA
ITAGVVEHLNKRISGVLPFVGGDRFSPFSTHGCKIVANVTINISWLNCTQV
NGERFVAIPCSVDYVAMLRVYSGKFLVMSRECYPDGLCYAMNRNLCAPCCOP
RESBYPGSGVSVARLKVILKRGDEALNIGRGYVSRARFACDYDKYASLMRL
AGYVGEVISHSTLASITVLPKGRADRPKISFGSCCELYVESVNDAPRLSE
KULRVPRNGVTESSADVGNVTRKRDARIDKRFHSLAGSVNKKKLAAGVLYR
VGGMDPHNSFLTOAGYHILYRKTSYSCVCELPDGLPRDVI PCSDVAMNMF
AAGRPVLVTRDKIKNGYCIYLAHCYASAPLKGRHVPDPTGANTPAKLRMY
SVIGDRSLNLYGFTSRGI FICDIDAVKDRMSALVAGKEMGVASDITPA
MKRTIYAVDRLYGSNDLKLSTIEKDLIDFNDVQSLKQRPVIRVPEYSEVQV
SLTFYQPELKFSTHSDHPAAASRLNEETLVLCGHVSVDIGGDLPLANKT
ORRVHCRPVLGDKAORVRELQYSSVSGDEKLEGPNDVCHPLGCHEHS
STMMVQVYASLCEICSMIKKKARITVLTVPGLDRECYMESLDCDIEYD
HAYVMYKFGSSCSYKSLTIKDMTPVYVIGGLPSVEMVEIRMGVAFKTKSEV
CPNISCGLRPRANSDDVYKLPREDKRRMLGPDITTIYDSKFSVRPYYVCN
CSAVNSKTFEWMVSPFKSKSRVITSGKTIHKDVNLDKYESFAVAMLSGVRSFLA
SEYLANKLHSGDSCSPIEATSPVLRBEINMTINFECLQVYKVAATVLSFLD
LDATIESITDFAECVAVDLBELGCLANANEKIRLADSMANLASFIVDIDAK
PVAHGHSPSEVEKEKEHAKGLRGSGKAGVLEPLDPLVSASRLFSCVDAQOQIK
SYVALTLVAVSLDONTENLSPIMRVISKGYSCMSFYLANRGLANRVASQVATVEI
AYLCANYSVSKATYVTPSAACVMAGMSPGSGDREYKSVHARTVOUTPDI FRED
SSYPLIBLSAICGAVITLPSGTSVSLAPLLOTKFSLILYRANVRYTHSST
TADGVRSVFSQITLILVNGTGNLVQIAPLGVPLLVKCVGLTFSMSDITSG
PIKGVSEFETITILYKRGQVLDLKAQVQVPSNKALFRDVTQEARISVIMQIP
TAVVDGLCAKASKYSGIVGKTNVGEKSGRGDGDHEEYFSCDESETPLRGVSGS
FSILFSPRLVARRARILIGIKRCPFKIPESISIKHELTPLFKGSDVAGVYFSL
SEVADALSIIEDEDEIBENLIGDPRMADSLDMNYPDSEDFLASVNGCAGLR
GCGRRVGMALINVYVNSLIMARAASNSTPTSLIFSRGLVDRIANFENESVEE
LVRYKYNFLSKRKGILGRNHYNSFGELVPLLANKLSRTIYAYVMSCKKY
CLASEREKFFALCVLTGLPYPARLVGGVDCKLSSLRAVDRIYFLKSCDA
MNGRLCFHFEDNGDSTAGLKGGSQTFWGLLRLVRLALSAGAGCCVSHDL
FELCVSLVEEYMAKILGICSPSRPISISSALFALQOPFVEYSGKSTTLPFLK
ISPRVPLIKGFEMVAGSPPLRGVRRLIAVKILRRKGLRTNGSPQORAPV
VIPENRYTNGCELSERMOGVEALRESELIIELEDAVVEKLNKRNRREADESER
PMHSTYTDVTSFCSKANSVGTALERAVLVEDAVKSEKISMDARDIVAKGTVDSEITN
TUSDDEALIEVPVSEBDDPKTRYSEYINRLNLSFEPKPIYVNDNKTGGLTNAY
REFYVQELAFELHSLCAITDOLRVNERSMAPCNDQGLYRNGSSTIQGQV
RSNTRKDFHDFLDGKISVNRRRSGNVLYHDLSPLSANLPLAGPERSVFNIS
SVLILYEPAGGKITTILIDSPKFKGVESEMTILTANKSQUEVILKKEVEDONI
ECORKDRPFPKSIITIDAVLHMRGCDADILFIDCFMVHAGSVLACIEFCHKY
MIFGDSROIHYIENELDKCYGLDFVDCGVNGVNSYRCPMDVCAULSVYVNM
IATVYSGESGSMRINEINSVDVLDPMGSTIYCMOSKLEISHFISKSLXN
LTVHAGETAYARNVRLKQDEPFPKSIPIHITVALSRHSDSLTYVLAARGDQTC
DAIOKABELVAKRFVFPFISRGSVTINDVKKDENDNRCAASAPUSVINDPLSVNP
GTATDFGDSLSDSTGTFEGGASGIVARNISSNLTDDHKQV"
<8901..10280
/note="53 kDa; translated via ribosomal frameshift"
/codon_start=1
/product="RNA-dependent RNA polymerase"
/protein_id="AAN63467.1"
/db_xref="GI:24636916"
/translating="SAVSQAIIPRRKSLQENLYSPARNYNTSTCDRTSMTSFEA
MAANCIRCFDLTFSSLRDVISITRSGIHEOMLEKTPSOIKALMNDISSPEIDDE
ICREFLAKYRDAKYLDSCLTKSPANQINIFKRSINAFISPIFNEVKRINCSCKP
NIKFFTEMTNRDPASVSNMGLDDVYHIGVDSPKSDQDAFVSEFMEYKEDV
DELLAIIMWCGERLIANTLGO.SFTINORSGASNTWINGSLVTLGLTSLYDVR
HFPAIYSGDSILFISRKISNYADICADMGPEPKSPSVFPCSKPYVMGCHTF
FVPDPYKLVKQAVKEDVSNDFLPEFTFSKDLTSDPNRERYIOKLAELVALKYDVQ
SGNTALALSVIHCRLSNPLSKLYPRVKGQVYITSVKALFPGSGSLYESVYTPG

```

---

```

CDS
HANNVNDDE"
10373..10543
/note="6 kDa"
/codon_start=1
/product="p6"
/protein_id="AAN63468.1"
/db_xref="GI:24636917"
/translating="MDVLOQBECLYLLNQAVFATFPVILLIRIVKSPRRNHEPTI
SAVSGQSPSTV"
10552..12351
/note="65 kDa"
/codon_start=1
/product="HSP70-like protein"
/protein_id="AAN63469.1"
/db_xref="GI:24636918"
/translating="MVYFGDLPFGTPTSTVCYVKGDKYSPKONNSAYIPVLYFSEPT
NHMFVGAESILMNAKVRGSEFFDLKRWGCDSSNFSYDTRLKPRYANRPYKIGG
LNDVILIGNYGVTRSEVHLPEGLIASPIKAMVSCAENAFSCGTGVICSPANDSVQ
RNFIDQCVSLSGYQCVYMWDEPSSAALASCNLANKMSANLAVDPFGGTDVSIISYR
NNTFVVRASGGDLNIGSDVDRFTLTLFSLTSLPEPLSDISILNKESSLTDAIYV
TLKSGDDKEBVRNKSILITVMPLPYNRITLINLEATLKYAKMGGLRSYKCDLVLI
GSSVPLGLEDILSKRRIDRIILKVADPRAAVNCALYSKCSGSGGLLVDCAHT
IATKMSCDQITICAPAPRIPFSGTTPYLIPKARNQROQIAYFEGEYVCPBRKTC
GSNTRFDTIGTVSYSPITFYLDTVITSSGALSFVVRGPGKVESLITGPAYNPSVY
VLGRSVREELIISLNKRVHLILHRRADRRLLEEGEAVRYETVEADVLEKFSY
NASSLPDEDEVEPLMGKSVQYLRGRLIEIPL"
12278..13882
/note="62 kDa"
/codon_start=1
/product="HSP90-like protein"
/protein_id="AAN63470.1"
/db_xref="GI:24636919"
/translating="MSHNSWGSILPKRYGGEADWERYLKSIVAHANIKTLPDVQLYG
GRYKYSFESATPSPFEOELGLMLSERDVGWKMCGITVEEAMYOITNBEAYPFA
EKCDDPYTEGQKTMEDVNMGRMSLNDIDKILIAOCMSLSCGGLINDPDEKFA
VALTRKNEIYVNDSSNNTIDCGVDGYVYASNSFNGCSQNSQANISILYEXYCAIYR
YIENTDLYFTSPDPIPLITGLIYPCCEYNIFFSYTRYKNDNPPFPLTNPLISDI
VFQWYKPAVDRLPLFDLSAALTVEPLSLVDSQVAVIYHLYRVESTADPAIDALE
EKDLAMKSSPHSLSTAQLWVGFCYGEFPAQRAVQRPQVATKTPNSVGGFETNMK
NVEEFDDIQCIEVYNSLRQPNGARHAEAFKVKKNISPKPISRLNIPREFYLVN
DYFRANRSGSEEEVLINNIISVDRVLCARACSTLPSAKRFSKHSKSVPSLRQ
KVYRPLIALNSLYEFHGRGRV"
13845..14516
/note="25 kDa"
/codon_start=1
/product="coat protein duplicate"
/protein_id="AAN63471.1"
/db_xref="GI:24636920"
/translating="MSGGTJAVGSESEALIEISGVILSTRKEAVDKFENELNENYSV
DSSRLSDQEVLEKSKESPKSELASDEHYVHILFPLRSALISISDVKYVGS
KYVYEGKTYTQDVWVPMIKGLIKOYRVEDVLAFCACAGDYLITVAPAMSEFKTK
AIGKGLVAGVEYLCADPLSGTSLKMSDHDAAVSIIVAANAVIDSAPFTGGRKIVSLY
DLGR"
14604..15200
/note="22 kDa"
/codon_start=1
/product="coat protein"
/protein_id="AAN63472.1"
/db_xref="GI:24636921"
/translating="MELSDNLSGLVITTDASSINGVDKLSLEVITKMLYOKAPSO
GIEYVGLLVALAARTTSPYQRADSDIIFONTYGDFTVAVTSGDLKVEGCAPLN
RFTNKLRTGFTTEAVYDPCVAYXKMPQUNAAAEADILPAEDSYLADFDGACPKS
ELQGRKMFASVYALKTEGVYNTPVSNRLQGRREV"
15200..15685
/note="19 kDa"
/codon_start=1
/product="p9"
/protein_id="AAN63473.1"
/db_xref="GI:24636922"
/translating="MEICQNPDEALVLTNTQNTLLVYKSDDEVNLEPKLIGCYLRY
SHRGDVTGQCNEDVYKDEGAHHVYIRERTVRYESAVKEYNADCVKFFLETGDV
RWFFIQSDIKRRAAHRLTFPEANPPFGSGHGMVEYCLKVLLITEVESVPSCEBRN
R"

```

CDS  
15687..16304  
/note="24 kDa"  
/codon\_start=1  
/product="p24"  
/protein\_id="AA063474.1"  
/db\_xref="GI:24636923"  
/translation="MRVIVPYEADILRSTEMRNIDSGALDTRCICAFSTLRD  
LHCARXSYQWCVDTGACORCAKRLIDTVESNLRAHPLIREKVAIHCXDEPELV  
APITRRVYLVGVAAAEAVKIMRILSKVLKLSLEMAFPYSPAPMAEMLDLRYT  
PLKIRYDLDDVDITNALCEDEDEIHVYVTKEDENHLEDEDC"

Query Match 62.7%; Score 374.6; DB 14; Length 16527;  
Best Local Similarity 76.7%; Pred. No. 1.2e-109;  
Matches 458; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

ORIGIN

QY 1 ATGAGCTTATGATGTCGACAGCACTTACGACCTGTGATACCGAGCGCTTACTCTA 60  
DB 14604 ATGAGCTTATGATGTCGACAGCACTTACGACCTGTGATACCGAGCGCTTACTCTA 14663  
QY 61 AATGCTGTGACAGCAAGCTTTATCTGCTGAAGTGAATAAATGTTGGTGCAGAAAGG 120  
DB 14664 AATGCTGTGATGATGAAGAACTGCTGCTGGGAGAGCTATATAAATGCTTGGCAGAAAGG 14723  
QY 121 GCTGCTTAACAGAGGATAGAAAGTGTGCTGCTACTCTTACGCACTCCGCGGCAAA 180  
DB 14724 GCGCCAGCCAGAAAGTATGAAACAGCTTTGGGCTACTGCTACGCTCTCGAGAGGAGA 14783  
QY 181 ACCAGCTCTCTAAGGTTGACGCGGAGATTACAGCTTATATTTCAATAGTTTCGGA 240  
DB 14784 ACCAGCTCTCTAAGGTTGACGCGGAGATTACAGCTTATATTTCAATAGTTTCGGA 14843  
QY 241 GAGAGGATGTGTGATGATACAGAGGTTGACCTTAAGAAGTACTCGAGCGGTGTGCGCT 300  
DB 14844 GAGAGGATGTGTGATGATACAGAGGTTGACCTTAAGAAGTACTCGAGCGGTGTGCGCACT 14903  
QY 301 CTGACTAGGTTCTACTAATACTAGAAAGTGTGCTGCTGCTTACTGAGGCTTACGTT 360  
DB 14904 CTGACTAGGTTCTACTAATACTAGAAAGTGTGCTGCTGCTTACTGAGGCTTACGTT 14963  
QY 361 GACTTTGATGCGGTATGAAGCAAAATTAACCCCACTCAAGCGCGCGGCGGATTTGGG 420  
DB 14964 GACTTTGATGCGGTATGAAGCAAAATTAACCCCACTCAAGCGCGCGGCGGATTTGGG 15023  
QY 421 ATTCCGCTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
DB 15024 ATTCCGCTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 15083  
QY 481 GAATTACAGCAAGTGAAGATGTTGCGAGTATGTAAGCTTAAAACTGAAGGTGA 540  
DB 15084 GAATTACAGCAAGTGAAGATGTTGCGAGTATGTAAGCTTAAAACTGAAGGTGA 15143  
QY 541 GTGCTAATATACCAAGTGAAGATGTTGCGAGTATGTAAGCTTAAAACTGAAGGTGA 597  
DB 15144 GTGCTAATATACCAAGTGAAGATGTTGCGAGTATGTAAGCTTAAAACTGAAGGTGA 15200

RESULT 6  
BYU51931 10545 bp RNA linear VRL 05-Apr-1999  
LOCUS DEFINITION Beet yellow stunt virus helicase gene, partial cds; and RNA  
genes, complete cds.  
ACCESSION U51931 L20761  
VERSION U51931.1 GI:1388128  
KEYWORDS Beet yellow stunt virus  
SOURCE Beet yellow stunt virus  
ORGANISM Beet yellow stunt virus  
VIRUSES; ssRNA positive-strand viruses, no DNA stage;  
Closteroviridae; Closterovirus.  
REFERENCE 1 (bases 2734 to 5576)  
Karasev,A.V., Nikolaeva,O.V., Koonin,E.V., Gumpf,D.J. and  
Garnsey,S.M.

TITLE  
JOURNAL Screening of the closterovirus genome by degenerate primer-mediated  
polymerase chain reaction  
MEDLINE J. Gen. Virol. 75 (Pt 6), 1415-1422 (1994)  
PUBMED 94267425  
AUTHORS 2 (bases 1 to 10545)  
8207405  
Karasev,A.V., Nikolaeva,O.V., Mushagian,A.R., Lee,R.F. and  
Dawson,M.O.  
TITLE Organization of the 3'-terminal half of beet yellow stunt virus  
genome and implications for the evolution of closteroviruses  
JOURNAL Virology 221 (1), 199-207 (1996)  
MEDLINE 96266429  
PUBMED 8661428  
AUTHORS Karasev,A.V., Nikolaeva,O.V., Lee,R.F., Wisler,G.C., Duffus,J.E.  
and Dawson,M.O.  
TITLE Characterization of the beet yellow stunt virus coat protein gene  
JOURNAL Phytopathology 88 (10), 1040-1045 (1998)  
PUBMED 94267425  
AUTHORS Karasev,A.V., Nikolaeva,O.V., Lee,R.F., Wisler,G.C., Duffus,J.E.  
and Dawson,M.O.  
TITLE Direct Submission  
JOURNAL Submitted (20-MAR-1996) Alexander V. Karasev, University of  
Florida, Citrus Research and Education Center, 700 Experiment  
Station Road, Lake Alfred, FL 33850-2299, USA  
On Jun 25, 1996 this sequence version replaced gi:507953.  
FEATURES  
source  
location/Qualifiers  
1..10545  
/organism="Beet yellow stunt virus"  
/mol\_type="genomic RNA"  
/specific\_host="Sonchus oleraceus L."  
/db\_xref="taxon:35290"  
<1..2133  
/function="replication"  
/note="ORF1a, similar to helicases of other  
closteroviruses"  
/codon\_start=1  
/product="helicase"  
/protein\_id="AAC55658.1"  
/db\_xref="GI:1388129"  
/translation="SSPLCAVMDLRSGYCFISNKYFSGVSDVYFKAVPTTINKRPP  
TLDSIFPVHVNHNVEVQKLFLEGGGFINENSGSGSVGEENDEFEARNDKMD  
RESGTSSTRANFVNDVABGKTEBRASGSDVSELYLOEPLFELINKLRPSSE  
TLVSKFLNLTENAFPEPMTEADLANSVREFFYLOEPLFELINKLRPSSE  
LVEFDRKICRAGDDARLLVRESDASLYGQDGRPLDMEYEVEFTGULPNDEK  
SGKRLPHANTKFIANSFELNSNSYRNFIEENDCRILYEAPGQKFTLLASFYK  
MEKKRILIVLANKSVEELIKINDSKREHETKTLKFAKARENYPASDSNVY  
TIDSVLNMHGTGKCDVLFVDECFVHAGAVTAVNTLVPSVFPVGRDTHIERNE  
YDVASFSDLDLVAAKRIYQVGYRCPMDVCGWLSXGYPTVATTVVSEKSSLLI  
TEINSVDVVASKRTYTLFLOSEKLEKRLAKRYKATYKTYHAGQPTIRDTYVY  
RTKQEDAPFSSFNHINVAITRHTESLTVAVLARNDNIAALICANELVDKRLTP  
HSFGQSVLNIIVEPVYTDNSCKASAPINSINDLEDDVPGSTLNFQDTSAMEQO  
PFESGANNVTVRDSAKPQSGTDHDEQV"

CDS  
2132..3514  
/function="replication"  
/note="ORF1b, similar to RdRp's of other closteroviruses;  
the BYSV RdRp is presumably expressed via a +1 ribosomal  
frameshift"  
/codon\_start=1  
/product="RNA replicase"  
/protein\_id="AAC55659.2"  
/db\_xref="GI:4559421"  
/translation="SVYSSOAIIPKRBPLOENLISYEARNVPTTCRFPSPGHEXRA  
MAANTVKGCFIDKRYAIRSDITAITGCGFPEMISKSPGSKSLASITLPLNLEE  
IVRFVLMYKRDARVLDASCLVQKIPPAOTLNFHRAKNAITSQCFDEKRVLCISL  
NIIFTEMTNRKQFADIASNLGAFENVETIGIDSKYDKSDQDAFKSDEALVSEFG  
GEPDELLVMMGEYRSEKNTMDQVGTTCVQCKRSGASNTWINSVTLGILMAYD  
VSKFQALFVSGDDSLIFSADEIANYABDIQELGSETFTLSVPVFCSKLVPFGDK  
CVFVDPYKLVKYGASGRSLSDLELFEVFSFDLTFEFGDERLVANTLSLVHLYKE  
FEMGNTKALSTIHLRNLFSFSLFVKRIGWVYVQKAKYILKKFLGYNIEPTTT  
FGDANFYVKE"  
3522..4343  
/function="unknown"  
/note="ORF2"

CDS  
 /codon\_start=1  
 /product="p30"  
 /protein\_id="AAC55660.1"  
 /db\_xref="GI:1388131"  
 /translation="MALYKSAALIVDPDRRTYSFSPFNRAIFSAACARVFOGVSVCS  
 GSVASTLPTLTSVFEVQNSGEPLSKRYRMNDERRITGAMTIPHSDAKIFVQ  
 VKQIDGTGGLLYNPHLIGGITVROFDLQGCENAAALACTPSRSPFASVIL  
 CDCPMSFTINTSDNNRSACTDAFYRRVFEVDSGLDITVDNSERNRFEVLRABE  
 SRSDPRVDVDSSEPEKSVLNRMCMCLFLLVLLAGDTIYVFNLSRFPVS"

CDS  
 /function="unknown"  
 /note="ORF3; similar to small hydrophobic proteins of  
 other closteroviruses"  
 /codon\_start=1  
 /product="p6"  
 /protein\_id="AAC55661.1"  
 /db\_xref="GI:1388132"  
 /translation="VDCILRAFLPFGALVCFPIAAAFAPFVKNTHSQTVDVI  
 RQEDLAG"

CDS  
 /function="unknown"  
 /note="ORF4; similar to HSP70 homologs of other  
 closteroviruses"  
 /codon\_start=1  
 /product="p66"  
 /protein\_id="AAC55662.1"  
 /db\_xref="GI:1388133"  
 /translation="WVVGVDVPGTTFSSVCVFNISGLHFKQNSAVITPCLFLYSDT  
 NMSFGYDAETASLDPNVKGFPFDLRWGCDSNTLEEKSLKPHYVSTLSPFGQ  
 SRKPTLGSYSGVQMSGLSLALP10ALVKSAAEPCECTELIVSPANCDMG  
 RLPTENCNLGFTCVHMMNPSAALASTCOTDMSANLIVDFOGGTVDSVLSL  
 NQTFVRSAGGDMNLGGHDVDFARAKIYOMANLPDEBADISLSKESLQIDVLY  
 TKDQSGKTVYVSRGLAVIVPVDRTIKVAKVVELYVNMQLKADAAASVIL  
 VGSVYELGKSLIOSVDVSECDLDPRAAVALGALVSSCSSESPLHLDGASH  
 NLSIPNYGESIVVYPAGAPVPTGTRDINLASCVSGSYSPVLPEDRCKCYNKVY  
 FSGIVPKDLGVGTDPRTITVTLATRVSSGVKPTGLSAAKRVGAGVGVDFRSK  
 ESVSIRSLITELHTDNRVLLATLTITREARQKFSYSEKQHLDSLVSGNLDREKSK  
 FNGVNOTADVRCKILGKSVQKTLRGARVELSYRNIYEVQTLIKI"

CDS  
 /function="unknown"  
 /note="ORF5; similar to analogously positioned proteins of  
 other closteroviruses"  
 /codon\_start=1  
 /product="p61"  
 /protein\_id="AAC55663.1"  
 /db\_xref="GI:1388134"  
 /translation="MSRRPTFAGYSMGLSEKRYGEBEMKSYLTETSMKYKPKLSESI  
 TFDSSLSLTAELRPARSGTALEYIALIFSDSTTKSEKLSRYGLNQNHSIIY  
 ATELEVTDVTKIGCKPTISAVSEPMGSRASAOVHRCMSLNSCGELINPNTAFI  
 OLVPDKAVTBOAONVNSGVSIVLYVCLQLYNDSKKKSNAGRTOLMESVSTIRPF  
 OHSIDYRSPIDNPLITGVLYDLCIENNVYSGYIKULNPRFKQIYPMIDIDFY  
 SWEIYAPBERLPIIDPEITKEVPTMSVIANVLSNKLVIDSTENNSILAEK  
 IISILCRNNEGIDSGALAAFCYGYTRDQVVRVPTVEDLGFSPKPYMSVVE  
 LFFDLQKRVDEKILRRFNAGAKGEMIVFKLGISPPITRLNAPSKSYINIDYF  
 KOANSLIGTEPEKIIICNIADVMMCAQRISYKAPLIAORGEALNSAKIRTLDP  
 TLVRALEKCLNOAPSMMWTLITNLR"

CDS  
 /function="unknown"  
 /note="ORF6; similar to analogous Byv coat proteins of  
 other closteroviruses"  
 /codon\_start=1  
 /product="p25"  
 /protein\_id="AAC55664.1"  
 /db\_xref="GI:1388135"  
 /translation="MPQPGAEIVHNANKSLLEVFSSETRKQKFPNNFDKTRQ  
 VNPINLNEDELREVGLKTELTKNLALDEDIIVAHVAFLLASVSTSPKYEYGS  
 YSISIDOKRYTNDAMTFPOVKILLASHNKNGELRACALAEQVLSVARTGDAFGT  
 RSVGRGAPSGSEYIGADFLITSTCPMNSDHRAYALASNALDRSAASIDIKMSVL  
 YDRGKVVIT"

misc\_feature  
 /note="transcription start site for coat protein  
 subgenomic RNA"  
 8617..9264

CDS  
 /note="CP; ORF7; similar to coat proteins of other  
 closteroviruses"  
 /codon\_start=1  
 /evidence="experimental"  
 /product="coat protein"  
 /protein\_id="AAC55665.1"  
 /db\_xref="GI:1388136"  
 /translation="MAGNDEGSDDSASQTKMDIPEPNEFARASATCLNGENK  
 KKLPERSVAKODVSEGIPTLITMTIYALATTSKGIQEDTLPVSKIDAVN  
 VTIYEDIKFNPASITLILNRYKRLKVPARETEEBELRVQKRLHPPIAANNGI  
 PADYSYLAADVPQTSNLKHEQVAVLEGNATATASGTTRESAVNLKYLGS5SK"  
 9261..9743  
 /function="unknown"  
 /note="ORF8"  
 /codon\_start=1  
 /product="p18"  
 /protein\_id="AAC55666.1"  
 /db\_xref="GI:1388137"  
 /translation="WNTYBSLQVTVVHERKALVITDODATIFEHLYGVHSSPYVTL  
 WYHGRGAMADSVISFTYKELQDVNLSLQYTRMLVRKTTSSSLTELINATNAV  
 SIFRCNAPTLIPVTSKGPESMDVAVLYNGIIGTPVDPLHLSIVSEYGNITPFAIE  
 " 9746..10306  
 /function="unknown"  
 /note="ORF9; similar to the 20-kDa protein of CTV and to  
 the 21-kDa protein of BYV, two other closteroviruses"  
 /codon\_start=1  
 /product="p22"  
 /protein\_id="AAC55667.1"  
 /db\_xref="GI:1388138"  
 /translation="MKLLSDSYDYDINLANKCLNELNKSVPLESCLASYNELLIK  
 VMAOSDADARKEKNEMLPDLNVMTDLDELVRDIPOLIAEINKENCSNDQV  
 KMIIRNVINSHNSRSPPTFEOIOSFMRGVKTSQNLNLYKKKFEVITVYFCNA  
 LIDPKMVSFKTVEVNLGENLESFL"

## CDS

## 3' UTR

## ORIGIN

Query Match 14.5%; Score 86.6; DB 14; Length 10545;  
 Best Local Similarity 50.4%; Pred. No. 3.2e-16;  
 Matches 212; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

QY	48	CGCCCTTAGCTTAATGCTGTCGACGAAAGCTTTATCTGCTGAAGTTGAAAAATGTT	107
DB	8721	CGCAGCTTGTTGATGTAATAAACAAGAAAGTTGTCGAGAAATTTCTGTGAGACT	8780
QY	108	GATGACGAAAGGCGCTCCTAACGAGGATATAGAGTGTTGCTACTCTCCTTTACGC	167
DB	8781	GAAAGCTGAGANTGCTACTGAATCCGCGCATCCACAGACTTTGGAGATGACGCTGAGC	8840
QY	168	ACTGCGGCAAGAACCGACGCTCTCCTAAGGTTCAGCGCGAGATTGACAGTTATATTTC	227
DB	8841	CCTAGCTACCTTAAGCACTTCTCTTAATAATGATATGAGATAGACACCGTTAGTGAG	8900
QY	228	AAATAGTTTCGAGAGAGAGATGTGATGATACAGAGGATGACTTAAGAGGTACTCGA	287
DB	8901	CGCTAAGTAGATGCGGTCAACGTAAGATPACTTACGAAGACATTAAGAACTTGTA	8960
QY	288	CGGATGAGCGCTGCTACTAGTTGCTACTAATAAATCTTAAGACTTGGCGGTACTTAC	347
DB	8961	CTCCTGACGTGCTTAGAATCTTAAGAACTTAAGAACTTCCGCTTGTGCTGCGACTTTGA	9020
QY	348	TGAGGCTTAGCTTACTTGTATGCGGTATTAAGCAAAATTAACCCCACTCAAGCGCGC	407
DB	9021	GAGAGAGTACCTTAGCTTGTGAGCAATCAAAACATCCCTCCCAATATGCAAGAGC	9080
QY	408	GAGCGAATGGGGATTCAGCTGAAGATTCGCTTACGTCGAGATTTCTGGGTACTTG	467
DB	9081	GAAACAACCGATATCCCGGTATTAATGATTAAGCAAGCTGACTGCTCAAACTAG	9140
QY	468	C 468	
DB	9141	C 9141	





mat\_peptide  
/gene="ORF1a/b"  
/product="papain-like leader peptide"  
/function="autocatalysis"  
/function="replication enhancer"  
1872..6233  
/note="ORF 1b; presumably expressed via a +1 ribosomal  
frameshift"  
9445..9609  
/function="unknown"  
/note="ORF 2; ps; small hydrophobic protein"  
/codon\_start=1  
/product="6.4 kDa protein"  
/protein\_id="AAC25116.1"  
/db\_xref="GI:3283078"  
/translation="MDCVLRSYLLAFGWICLFLCLVAVFVYKQILFNTPEPSN  
EARFNSTVY"  
9609..11405  
/function="unknown"  
/note="homolog of HSP70 proteins; ORF 3"  
/codon\_start=1  
/product="65-kDa protein"  
/protein\_id="AAC25117.1"  
/db\_xref="GI:3283079"  
/translation="MVFGLDPTGTSVAVYGERLYFKORDAVIPFVFLHSDT  
QVAFGIDAVYLVNDSPVRCGPTVDLKRMGCEENVGYLTKKHVITTELLKVAS  
SKVTRIDCTSGVLPQNALPGLATFVRLISTASBARCQCTGTCVPPANNLO  
RSFTSCVNLGSPCYVWNEPSAASASRIKATPVLVDYDFGGFVSVSAL  
NNTFVVASGDMNLGDRIDKAFVHLVKKQLPNYKIDISFLKESJKVSVLNF  
PVSEONKVDLVNSELAEVAPFVEERTIKIVKEYEKRSNLEBSVAKLMTV  
GGSSYLEGLSRVSVFVECLVLPDARAVAGCALVSACLNDSFVCLVCAHN  
LSISKYCESIVCPAGTPIPTGVTVNMTGNASVYSAALFEDGPMVLCNKRIF  
SGVYALGVNVSQNTPTVPLTEIVNSVGTITPVGTVKILVGNAAVDESSY  
QLGRVAVADJHKNSDKVLIHATITPFRKCLTEGDALFLKRLSADYRBAKFS  
SYDVAVNSSELLGRVLPKILRGSREKLDV"  
11311..12972  
/function="unknown"  
/note="ORF 4"  
/codon\_start=1  
/product="64 kDa protein"  
/protein\_id="AAC25118.1"  
/db\_xref="GI:3283080"

mat\_peptide  
/gene="ORF1a/b"  
/product="putative methyltransferase and RNA helicase"  
/note="ORF 1a"  
/codon\_start=1  
/product="7983,8045..9378)  
/gene="ORF1a/b"  
/product="RNA-dependent RNA polymerase"  
/note="ORF 1b; presumably expressed via a +1 ribosomal  
frameshift"  
9445..9609  
/function="unknown"  
/note="ORF 2; ps; small hydrophobic protein"  
/codon\_start=1  
/product="6.4 kDa protein"  
/protein\_id="AAC25116.1"  
/db\_xref="GI:3283078"  
/translation="MDCVLRSYLLAFGWICLFLCLVAVFVYKQILFNTPEPSN  
EARFNSTVY"  
9609..11405  
/function="unknown"  
/note="homolog of HSP70 proteins; ORF 3"  
/codon\_start=1  
/product="65-kDa protein"  
/protein\_id="AAC25117.1"  
/db\_xref="GI:3283079"  
/translation="MVFGLDPTGTSVAVYGERLYFKORDAVIPFVFLHSDT  
QVAFGIDAVYLVNDSPVRCGPTVDLKRMGCEENVGYLTKKHVITTELLKVAS  
SKVTRIDCTSGVLPQNALPGLATFVRLISTASBARCQCTGTCVPPANNLO  
RSFTSCVNLGSPCYVWNEPSAASASRIKATPVLVDYDFGGFVSVSAL  
NNTFVVASGDMNLGDRIDKAFVHLVKKQLPNYKIDISFLKESJKVSVLNF  
PVSEONKVDLVNSELAEVAPFVEERTIKIVKEYEKRSNLEBSVAKLMTV  
GGSSYLEGLSRVSVFVECLVLPDARAVAGCALVSACLNDSFVCLVCAHN  
LSISKYCESIVCPAGTPIPTGVTVNMTGNASVYSAALFEDGPMVLCNKRIF  
SGVYALGVNVSQNTPTVPLTEIVNSVGTITPVGTVKILVGNAAVDESSY  
QLGRVAVADJHKNSDKVLIHATITPFRKCLTEGDALFLKRLSADYRBAKFS  
SYDVAVNSSELLGRVLPKILRGSREKLDV"  
11311..12972  
/function="unknown"  
/note="ORF 4"  
/codon\_start=1  
/product="64 kDa protein"  
/protein\_id="AAC25118.1"  
/db\_xref="GI:3283080"

CDS  
/translation="MTTFSTPANYMGELFRFFGGRKMLMSEANASRPVYDRD  
FRSNGEVSRLKVFSDTGESFYREBSLITLTPKAEVCLGVAMEGLKGNRLSD  
YNSLNAVVDYRKVFGCKPNIQYTERVRLNGLAEPIVHEWOMLSNSGELINRDK  
TKRFVSLRFGKQVSTDEAVSSYLDYLSHCLNLYXTCNLSNSGKALYDELK  
YVLYVLETSLEFSLSDNPLVAGVYDMCFEINTLKSIFYLKNRSLYESTEDSN  
SEIFSMWQEPADVRSLDELDTTELLKIPINTHDSFFLYKNRSLYESTEDSN  
ELIKVUDSLITDNDDELKACQWVGFCHYGVFRACQKRVDRAEYLPEALGKT  
INSGVEEPDELQKMPSVSRFGCSLSHAFSIFRFGVFPPIRLNAVYYS  
YLVVDYRHKRAGLQDELITLNIIEPIVAEWCCREVALQARRAQRGEKFGQWG  
VKEVSPHASSIRYKNNESLNTVMKQVAGARQRLNPLRHKH"  
12920..13570  
/note="24 kDa protein; ORF 5"  
/codon\_start=1  
/product="minor coat protein"  
/protein\_id="AAC25119.1"  
/db\_xref="GI:3283081"  
/translation="MLAREGRDLIHFTENTRDMETFPNSYDLAEVSEVNPKNLRK  
ETBBLVITERKSELVITDEPVGHLARALRAANTITSTVNVVAGVAYTGKTK  
FLVDAVAFPLIECKMKFKKPVFPFCATFEDATVITARSIPKFLNRTIGKRP  
SGVEFLADFLITVCLNDHEKAIVLQASRAIDRAVSSVDKIVSLFLGRLS"  
13641..14255  
/note="22 kDa protein; ORF 6"  
/codon\_start=1  
/product="major coat protein"  
/protein\_id="AAC25120.1"  
/db\_xref="GI:3283082"  
/translation="MGSABPISAIATFENVSLVDQCLHGECDDKRRNFECLKLG  
VPEBKGLALGLCHYSCATIGTGNKSVQPTSPFICASSGSEGLFTHLELSFDS  
OKLEGGKRLRCGRTFQDYISFAEYGRAPLAPARHGLPABDHYLADFLST  
STETDLOCGRILLARNAHTPESSSPITSKQGRGLATOK"  
14252..14785  
/function="unknown"  
/note="ORF 7"  
/codon\_start=1  
/product="20 kDa protein"  
/protein\_id="AAC25121.1"  
/db\_xref="GI:3283083"  
/translation="MTGSVRLAOKRPFRVILLKGFVYVIAITEESEPVELPLY  
LHLELVNKKGLIESYIYIPKSCMTLRKSSVSYTRVSGKSEDSSESSYSGKTPDS  
KILNRKVTFTFESGIQVFPWYGRDQCVSSBYLWENVVGAHGTLPCLNCELDK  
SGELRLITFSPKNEVFL"  
14769..15302  
/function="replication enhancer"  
/note="ORF 8"  
/codon\_start=1  
/product="21 kDa protein"  
/protein\_id="AAC25122.1"  
/db\_xref="GI:3283084"  
/translation="MKFFPNDCTSRASISSESLRRVKEIGTNSPQSEVSECTENFN  
ELARFNLVTVHERREMEKHPKQSSSLRTPSPRLGEMIKETRAFLKRVVTPMKETA  
SETLNAFLFEYCRITGLTRDALARAKORRVSTVLFHSESLKEVTENNFSFTELLK  
LNLSTRVISOGLIAIV"  
ORIGIN  
Query Match 11.0%; Score 65.6; DB 14; Length 15468;  
Best Local Similarity 46.9%; Pred. No. 2.2e-09;  
Matches 240; Conservative 0; Mismatches 269; Indels 3; Gaps 1;  
QY 26 TTAGCACTGTGTGATTAACGACGCTCTAGTAAATGGTGTGACACAGAAGCTTTAT 85  
DB 13678 TTGAAGAGGTGAGTCTGTAGCCAAAGCTTTTGCACGCTGGAAGACCTGCACAACTAC 13737  
QY 86 CTGCTGAAGTGAAGAAAATGTTGTGACGAAAGGGGCTCTAACGAGGATTAAGAGTGG 145  
DB 13738 GAGGAATTTGGAAGAGTGTGAATTAAGAGGGGTTCCGGAAGATTAACCTGCTCGC 13797  
QY 146 TTTTGGGTCTACTCTTTACGACCTCGGGGGAAGAACACAGCTTCTCTAGGTTACGGCG 205  
DB 13798 CGTTAGGACTTTGTTTGTACTCTGCTGCGACGATAGTACTTAATAAAGTTAGCGTCC 13857  
QY 206 CAGATTACAGAGTTATATTTTCAATAGTTTGC--GAGAGAGGATGTGGTAGTACAG 262  
DB 13858 AACGACGCTACTTCAATCAAGCTTGTTGGTAGTGGAGAGGAATGTTCTCTCACTC 13917







## CDS

/db\_xref="taxon:12161"  
 1..612  
 /codon\_start=1  
 /product="coat protein"  
 /protein\_id="AAB17001.1"  
 /db\_xref="GI:161940"  
 /translation="MGSAPISAIATFENVSLVDQCLHGEDCDKLRNFECKLKG  
 VPEDKLGLALGCTGSKNKSVOPTSTFIASFGSKELFLNHGELRSFLDS  
 OKLBGPKHRCFCRFPKDYISPAKEYRGLEPPLARNRGLPMDHYLAADPISIT  
 STRLTLQGRLLIARENATHRESDSVTLKQGRRLANG"

## ORIGIN

Query Match 10.7%; Score 64; DB 14; Length 612;  
 Best Local Similarity 46.7%; Pred. No. 4.4e-09;  
 Matches 239; Conservative 0; Mismatches 270; Indels 3; Gaps 1;

QY 26 TTGCAACCTCGTGATPACCGAGCCTCTAGTCTAATGTCGACAGAAAGCTTTAT 85  
 DB 38 TTGAAAACGTAAGTCTCGTAGACCAACGTGTTGACGGTGAAGCTCGCAAACTAC 97  
 QY 86 CTGCTGAAGTTGAAAAAATGTTGTCGAGAAAGGGCTCTTAACGAGGTATGAAGTG 145  
 DB 98 GGAAGAAATTCGAAGAGTGTGAAATGAAAGGGGTCGGAAGATAACTCGGTCTCG 157  
 QY 146 TGTTCGCTACTCTCTTACGCACTGGCGGCAAGAACAGTCTCTAAGCTTCACGGG 205  
 DB 158 CGTTAGGACTTTTGTGTACTCTGTGCGACATAGGATCTTAAATGAAGTACGCTCC 217  
 QY 206 CAGATTGACAGCTTATATTTTCAATAGTTTCG---GAGAGAGAAATGTTAGTAAAC 262  
 DB 218 AACCGAGTCTACTTTCATCAAGCTTCGTCGGTAGTGGAAGAAATGTTCTCAGTC 277  
 QY 263 AGGGTGACCTTAAAGAGTACTGACGGGTGGGCGCTCTCACTAGTTCACATAAAC 322  
 DB 278 ACAGCGAAGTGAAGTCTCTTCTGACTCTCAGAACTTTAGAGGAAAGCCCTAACAAAT 337  
 QY 323 TTGAAAGCTGCGGCGTACTTTCAGTGAAGCTTACGTTGATCGGTATAAGC 382  
 DB 338 TGGCGTCTTCTGCGGACCTTTTCAGAGGATTAATATCTTCGCGGAGGATACCGAG 397  
 QY 383 ACAATATACCCCACTCAACGCCGCGGCGGAATTGGGATTCAGCTGAAGATTCGACT 442  
 DB 398 GAGAGACTGCTCCGATGCTAGAGCCACCGTCAAGGTCTACCTGCTGAAGATCAGTACT 457  
 QY 443 TAGCTGAGATTTTGGGTAATCTGCCGAAGCTCTGAATTAACGAAGTAGAAGA 502  
 DB 458 TAGCTGATTTATATCGACATCAACGAACCTTACCGACTACCAAGGTCTGCTGC 517  
 QY 503 TGTTCGAGATATGATGCTCTTAAAAAAGTGA 534  
 DB 518 TGTTCGCGCGGAAAAACGCCACCTCACACAGA 549

RESULT 11  
 CIBV3PH 6746 bp RNA linear VRL 13-JUN-2002  
 LOCUS Beet yellow virus genome 3', proximal half, for capsid protein  
 DEFINITION HSP70 related protein, RNA-dependent RNA polymerase and ORFs.  
 X53462  
 ACCESSION X53462.1 GI:58878  
 VERSION capsid protein; heat shock protein 70 homologue; RNA polymerase;  
 KEYWORDS RNA-dependent RNA polymerase; unidentified reading frame.  
 SOURCE Beet yellow virus  
 ORGANISM Beet yellow virus  
 Viruses; ssRNA positive-strand viruses, no DNA stage;  
 Closteroviridae; Closterovirus.  
 1 (bases 1 to 6746)  
 Agranovsky,A.A., Boyko,V.P., Karasev,A.V., Lunina,N.A., Koonin,E.V.  
 and Dolja,V.V.  
 Nucleotide sequence of the 3'-terminal half of beet yellows  
 closterovirus RNA genome: unique arrangement of eight virus genes  
 JOURNAL J. Gen. Virol. 72 (Pt 1), 15-23 (1991)  
 MEDLINE 91116305

PUBMED  
 REFERENCE  
 AUTHORS

## TITLE

JOURNAL  
 MEDLINE  
 PUBMED

## REFERENCE

## AUTHORS

## TITLE

JOURNAL  
 MEDLINE  
 PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## FEATURES

## source

## CDS

1990061  
 2 (bases 1 to 6746)  
 Agranovsky,A.A., Boyko,V.P., Karasev,A.V., Koonin,E.V. and  
 Dolja,V.V.  
 Putative 65 kDa protein of beet yellow closterovirus is a  
 homologue of HSP70 heat shock proteins  
 J. Mol. Biol. 217 (4), 603-610 (1991)  
 91171285  
 2005613  
 3 (bases 1 to 6746)  
 Rensing,S.A. and Maier,U.G.  
 Phylogenetic analysis of the stress-70 protein family  
 U. Mol. Evol. 39 (1), 80-86 (1994)  
 94343547  
 7545947  
 4 (bases 1 to 2677)  
 Agranovsky,A.A.  
 Direct Submission  
 Submitted (22-MAY-1990) Agranovsky A.A., A.N. Belozersky  
 Laboratory, Moscow State University, Moscow 119899, USSR  
 5 (bases 2678 to 6746)  
 Agranovsky,A.A.  
 Direct Submission  
 Submitted (27-JUL-1990) Agranovsky A.A., A.N. Belozersky  
 Laboratory, Moscow State University, Moscow 119899, USSR

## CDS

Location/Qualifiers  
 1..6746  
 /organism="Beet yellows virus"  
 /mol\_type="genomic RNA"  
 /db\_xref="taxon:12161"  
 <1..659  
 /codon\_start=3  
 /product="RNA polymerase"  
 /protein\_id="CA37549.1"  
 /db\_xref="GI:588947"  
 /db\_xref="GI:588947"  
 /db\_xref="GOA:Q66109"  
 /db\_xref="SPRMBL:Q66109"  
 /translation="DNORSGASNTWIGNSJETLGIISMEYNTNFKALFVSGDLSI  
 FSESPIRNSADAMCTELGFEFTFLTPVPYCSKFVMTGHDVFPVDPYKLVKGA  
 SKDEVDDEFLEFVFTSPFDLTKDVIDEIVILFLHVAHVKRGYSGDTVAALCHICI  
 RSNPSFKLYPKXGKGVHGXKAKFLAKRANCFRFPDPAFERFTLLTKLETVL  
 710..874  
 /note="unnamed protein product; ORF2"  
 /codon\_start=1  
 /protein\_id="CA37550.1"  
 /db\_xref="GI:58879"  
 /db\_xref="SPRMBL:008542"  
 /translation="WDCVLRSLYLLAFGLCLFLCLVPIWIMVYKQLFRTAQS  
 EARHNSHTV"  
 874..2670  
 /codon\_start=1  
 /product="HSP70 homologue"  
 /protein\_id="CA37551.1"  
 /db\_xref="GI:58880"  
 /db\_xref="GOA:P37092"  
 /db\_xref="SWISS-PROT:P37092"  
 /translation="WVVGFGDPTGTFSSVCAVYGEELYLFRQDSATIPYVFAHSD  
 QEVAFGYDAEVLNDLSVRGGEYRDLKSWICDSENYVDYELKPKHYTELKKAOS  
 SKSTVLDKSYGVTONATLPLGLATFVKALISASEAFKQCGVIGSVANVCLO  
 RSEYKLVDCYGVPCVYVWNEPSSAALSACRIGAKTSPVLYDGGTGPVYSIAL  
 NNTFVVRAGGDMNLGARDIDKAFVHLYNKAOLPVNYKIDISFLKESLSKVSFLNF  
 PVYSGQYRVYVYVNSLAVAAFPVETRIKIYVEVEKXCSMRLEPNVAKILMY  
 GGSSTLPGLHLSLSTIPVDECVLPDRAAIVAGCAIYSCAINDSPMLVDCAAH  
 LSTISKYESIVCPVAGSPFIPFTGRTVNTGNSISFLVGPVGVKILGNAAYDPSSY  
 QLGERVVAIDLHNSDKYKLIALTLVQFPQKRLDQKALFLKRLTLADYREARKFS  
 SYDDAVLNSSELLGRILPKILRSGRYEKLDV"  
 2576..4237  
 /note="unnamed protein product; ORF4"  
 /codon\_start=1  
 /protein\_id="CA37552.1"  
 /db\_xref="GI:58881"

## CDS

1990061  
 2 (bases 1 to 6746)  
 Agranovsky,A.A., Boyko,V.P., Karasev,A.V., Koonin,E.V. and  
 Dolja,V.V.  
 Putative 65 kDa protein of beet yellow closterovirus is a  
 homologue of HSP70 heat shock proteins  
 J. Mol. Biol. 217 (4), 603-610 (1991)  
 91171285  
 2005613  
 3 (bases 1 to 6746)  
 Rensing,S.A. and Maier,U.G.  
 Phylogenetic analysis of the stress-70 protein family  
 U. Mol. Evol. 39 (1), 80-86 (1994)  
 94343547  
 7545947  
 4 (bases 1 to 2677)  
 Agranovsky,A.A.  
 Direct Submission  
 Submitted (22-MAY-1990) Agranovsky A.A., A.N. Belozersky  
 Laboratory, Moscow State University, Moscow 119899, USSR  
 5 (bases 2678 to 6746)  
 Agranovsky,A.A.  
 Direct Submission  
 Submitted (27-JUL-1990) Agranovsky A.A., A.N. Belozersky  
 Laboratory, Moscow State University, Moscow 119899, USSR

```

/db_xref="SPTREMBL:Q06110"
/translation="MTFRSTPTANYWGELEFRFFGGGEMKLNLSSEASVSRPYSID
PFFSGVILSRKTFGESEFVREFSLILTFPKTYECKLQVAMBLANGNRISD
YVSESNIVLDTVGCENIOGTFVFKKINGNAVPLVAGLNSGCELINPDS
TRFVSLIFKQDLAESEDEIVSSSYDYLSCINLETNLSNKGKSLYDEFLK
HYDYLNSDLEYSRSDNPLVAGILYMCCEYNTLKSTYKNIKESFDCFLSYLPL
SEFPMNMRPADVRLFEIDAEILLKPTINMHDSITLTKNKLKYSPEDESDN
ELIKYVDSILTRNDDELKLAQWYCHYCYVGTOTRQTRVXRDAVYKLPALGPRV
YMSGVFEFEELQKMSISYRRAFCOSLSHEASVYKRGVCFPITRLNTPVKS
YLVNDYVRHVRKVGILQDELITLSNIEFVAMCCERLVALQARRARREPRQMG
TKNEISPHARRSIRVYKNDLSLNLMLDVAGRSORRLPLNRH"
4185..4835
/feature="unlabeled protein product; ORF5"
/codon_start=1
/protein_id="CAA37553.1"
/db_xref="GI:58882"
/db_xref="GOA:Q08538"
/db_xref="SPTREMBL:Q08538"
/translation="MLAPARGDLHFTENTRDAMETFPNSYDIAEYSEVNPKNLKNK
ETDELGVIRERKSELVITDEDFVYKALFALIRANNTTCKVNVGAYEYTGKPK
FLVQAMVPLIKCKCKENKPNPRFCATFEDAYIVIASLPLKLNRTIGRGIP
SGVEFGADPLFATSVCLNDBKALIVLQASRAIDRAVSSVDKIVSLFPLGRIS"
4906..5520
/codon_start=1
/product="capsid protein"
/protein_id="CAA37554.1"
/db_xref="GI:58883"
/db_xref="GOA:Q08531"
/db_xref="SPTREMBL:Q08531"
/translation="MGSAPISALATFENVSLADOTCHGEDCKLRNFECKLKG
VEDNGLALGCLVSCATIGTSNVOTPSFKAFGGKXELVTHGNSFLDS
OKLIEKPKIKRCFQRTPOKDYISLRKCYRQKPLPILANVHGIPADHYIADPFI
SEITLQOSRILLARENATHREPSSEPVSLKAGLGCTGR"
5317..6039
/feature="unlabeled protein product; ORF7"
/codon_start=1
/protein_id="CAA37555.1"
/db_xref="GI:58884"
/db_xref="SPTREMBL:Q08544"
/translation="MTSSVELAQCKPLRVLILKGFVYVAVFEBSSEHPLVLY
LHDFINIKRGKIASVYDPMSCMTRLKPSVSSTRASKESDSISPSGSTRS
KVLNRVITFENGVDVFGMYGLQRCVSSDYLTFENVFGAGCGLITLNCESLQK
SGEELITFSPKREVLKRW"
6034..6567
/feature="unlabeled protein product; ORF8"
/codon_start=1
/protein_id="CAA37556.1"
/db_xref="GI:58885"
/db_xref="SPTREMBL:Q08545"
/translation="MKFPLKDETSRALSRSLSLRVKEICTNSQOSEISECYDEN
ELASFNHLYVTEHREMEQHPNQSSKLVRBRIGEMKEIRAFILKRVVTPMKETA
SDLNAFLSEYCRITGLAREDLAREKRVKVSVLFFHSELKEVTEVNMFSYELK
LNLISVSISSQILGMAI"

```

## ORIGIN

Query Match 10.7%; Score 63.8; DB 14; Length 6746;  
 Best local Similarity 48.1%; Pred. No. 7.4e-09;  
 Matches 213; Conservative 0; Mismatches 227; Indels 3; Gaps 1;

```

QY 95 TTTAAAAATTTGGTGCAGAAAGGGCTCTCTTAAGAGGCTATAGAGTGTGTTGGTC 154
Db 5012 TCGAAGAGTGTGTAATTAAGGGGTTCCGAAAGATACCTCGAATCGCTTAGAGAC 5071
QY 155 TACTCTTTACGACCTGCGGCAAGACACGCTCTCTTAAGGTTCA---GGCGCAGATT 211
Db 5072 TTTGTTGATTTCTGTGCTAGCAATGACACTTCCAAAGATTAACCTCCACCGACGT 5111
QY 212 CAGACGTTATTTTCAAAATAGTTTGGAGAAGAGAGATGTGTAAGTAAAGAGGTGACC 271
Db 5132 CTACCTTCACAAAGTGTGTTGGTGGGAGGAGAACTGACCTCACTCAGCGTGAT 5191
QY 272 TTAAGAAGTACTCAGCGGTGCGCCTCTCACTAGTTTCACTAATAAATTGAACGT 331

```

```

Db 5192 TGAATCTCTTCTGGGGTCTCAAAAACTTTTGCAGGGAAAACTTAACAATTCCGGTGT 5251
QY 332 TCGGTGACTTCTCACTAGGCTTACGTTGACCTTTGTATCGGTATPAGCAAAATTAC 391
Db 5252 TGTGGCGTACTTTTCAAGAGACTACATATCTTCCGCAAGAAATCCGAGGAAATTAC 5311
QY 392 CCCAAGTCAAGCGCGCGCGGGAATTTGGGAATTCAGCTGAAGATTTGTTAGTGCAG 451
Db 5312 CTCCAGTTCCAGAGCTTACCGCTACCGCTTACCCGCTTAAGATCACTTACCGCGCT 5371
QY 452 ATTTCTGGGTAAGTTCGCCGGAAGCTTCTGTAATTACGAAAGTGAAGATGTTCCGA 511
Db 5372 ACTTCATATCGACGTCGACGGAAGCTCACTGACCAACAAAGTCGTCTGTAGCGC 5431
QY 512 GATATAGCGCTTAAACTGA 534
Db 5432 GCGAAACGCCACTACACGGA 5454

```

```

RESULT 12
LOCUS BYVUA
DEFINITION Beet yellows virus (isolate Ukrainian BYV-U) complete genome.
ACCESSION X73476
VERSION X73476.1 GI:405624
KEYWORDS coat protein; heat shock 70-related protein; heat shock 90-related
polymetase; RNA helicase.
SOURCE Beet yellows virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage;
Closteroviridae; Closterovirus.
REFERENCE 1
AUTHORS Agratovsky, A.A., Koonin, E.V., Boyko, V.P., Malais, E., Frotschl, R.,
Lunna, N.A. and Atabekov, U.G.
TITLE Beet yellows closterovirus: complete genome structure and
identification of a leader papain-like thiol protease
JOURNAL Virology 198 (1), 311-324 (1994)
MEDLINE 94082466
PUBMED 8259665
REFERENCE 2 (bases 1 to 15480)
AUTHORS Agratovsky, A.A.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-1993) A.A. Argatovsky, A.N. Belozersky Institute,
Moscow State University, 119899 Moscow, Russia
COMMENT Related sequence: X53462.
FEATURES
source
1..15480
/organism="Beet yellows virus"
/mol_type="genomic RNA"
/isolate="Ukrainian (BYV-U)"
/db_xref="taxon:12161"
/clone="BYV 936, x19, R9, R3, 36A, R2, 36, p43, 111-7,
112, 115, 1213, 1210, 124, 1313, 1311, 142, 143, 1311,
1311, 142, 143, 156, 1510, 154, 1518, 1520, 169, 1615, 175"
1..107
join(108..7983, 8045..9393)
/genes="ORF1a/b"
join(108..7983, 8045..9393)
/genes="ORF1a/b"
/codon_start=1
/product="fusion protein of papin-like protease,
methyltransferase, RNA helicase"
/protein_id="CAA51871.1"
/db_xref="GI:809539"
/db_xref="GOA:Q08534"
/db_xref="SPTREMBL:Q08534"
/translation="MAFLNVAIVSCAPAPAPHAGASIVDPSPCVPRYDDISH
FRLLTSLDFSPRPLSLANRVLHVSASTDNPLSLPLGFAHETVLELNGSAFSLBS
RHIDVNRPSVSPFTEVLVSLRTSRLFALCFELVCSKRGPCVEIASFTSP
CLVSCVAOITFAEMESIRFPTTLPAQFLFHKKXTPRPETLIHESGALTS
ALGVTSKNSRPITVKSAGKYEAVSISKDERSRRQCTRVSHKRLINKVE
PFFPEBPKDKRRASLPTEBQGITFGTLRPLSPTKPEBRLPKPREVEIPVVK

```



3'UTR  
 INLSLRVISQILGWAI"  
 15299..15480

Query Match 10.7%; Score 63.8; DB 14; Length 15480;  
 Best Local Similarity 48.1%; Pred. No. 8.4e-09;  
 Matches 213; Conservative 0; Mismatches 227; Indels 3; Gaps 1;

QY 95 TTGAAAAATGTTGGTGCAGAAAGGGGCTCTAACGAGGGTATAGAAAGTGTGTTGGTC 154  
 DB 13746 TCGAAGAGTGTGAAATTAAGAGGGGTTCCGAAAGATTAACCTCGAATCGCGTTAGAC 13805  
 QY 155 TACTCCTTACGACTCGCGGCAAGACAGCTCTCCCAAGTTCA---GGCCCAAT 211  
 DB 13806 TTGTTGTTATCTCTGTGTACATAGGAGCTTCAACAAAGTTAAGCTCCACCGAGT 13865  
 QY 212 CAGAGCTATATTTCAATAGTTTCGAGAGAGATGTGTATACAGAGGGTACC 271  
 DB 13866 CTACCTTCATCAAGCTTGTGTGTGGGGAAGAACTGACCTCACTACAGGTGAAT 13925  
 QY 272 TTAAGAAAGTACTCGAGCGGTGTGGCCCTCCTCAGTTTCAATTAATTTAGACGT 331  
 DB 13926 TGAATTCCTTCTGGGGTCTCAAAAATTGGAGGGAAACCTAACAAATTCGGTGT 13985  
 QY 332 TCGGTGTACTTTCAGTACGCTTACGTTGACTTTTATCGCGTATPAGACAAATTAC 391  
 DB 13986 TCTGCCGTACTTTTCAGAAAGACTACATATCTTTCGCAAGAAATACGAGGGAAATTAC 14045  
 QY 392 CCCAATCAAGCGCGCGGGAATTTGGGATTCAGAGCTGAAGATTGTAAGTTAGCTGAG 451  
 DB 14046 CTCGAGTTCAGAGAGTAACTGACGCTACGCTACCGCTGAGAGATCACTACTTACCGCTG 14105  
 QY 452 ATTTCTGGGATCTTCCCGAAGCTCTGTGAATTAACAGCAAGTAGAAGATGTTCCGA 511  
 DB 14106 ACTTATATCGACGTCACGAGAACTACTGACTCAACAAAGTCTGCTGTATAGCG 14165  
 QY 512 GTATGAGCGCTTAATAACTGA 534  
 DB 14166 GCGAAAGCGCACTCACAGGAA 14188

RESULT 13  
 LOCUS A41914 783 bp DNA linear PAT 05-MAR-1997  
 DEFINITION Sequence 1 from Patent WO9428147.  
 ACCESSION A41914  
 VERSION A41914.1 GI:2297469  
 KEYWORDS  
 ORGANISM  
 SOURCE  
 ORGANISM  
 Beet yellow virus  
 Beet yellow virus  
 Viruses; ssRNA positive-strand viruses, no DNA stage;  
 Closteroviridae; Closterovirus.

REFERENCE  
 AUTHORS Bojzen, K. and Brunsfeld, J.  
 TITLE DNA SEQUENCE COMPRISING AT LEAST TWO COAT PROTEIN GENES  
 JOURNAL Patent: WO 9428147-A 1 08-DEC-1994;  
 COMMENT Sandoz AG (AT)  
 FEATURES  
 Other publication AU 7122694 941220.  
 location/Qualifiers  
 source 1..783  
 /organism="Beet yellow virus"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:12161"

## ORIGIN

Query Match 10.6%; Score 63; DB 6; Length 783;  
 Best Local Similarity 46.6%; Pred. No. 9.7e-09;  
 Matches 228; Conservative 0; Mismatches 270; Indels 3; Gaps 1;

QY 26 TTAGCAACTGTGTATACCGCAGCGCTAGCTTAATGTGTGACAGAAAGCTTTAT 85  
 DB 126 TTGAAGACGTGAGTCTCGACACCAAGCTGTTTTCACGAGTGAAGACTCGCAAACTAC 185  
 QY 86 CTGCTGAAGTGAATAATGTTGTGTCAGAAAGGGGCTCTTACGAGGGTATAGAACTGG 145

DB 186 GGAAGAAATTCGAAGAGTGTGAAATTGAAGAGGGGTTCCGAAAGCAAACTCGTCTCG 245

QY 146 TGTTCGGTCTACTCTTTCAGCACTCGCGGCAAGAACCAAGCTCTCTAAGTTCAAGCGG 205

DB 246 CGTTAGAACTTGTGTTGATTCCTGTGCGAGCAATAGTACTTAAATTAAGTATGCTC 305

QY 206 CAGATTCAAGCGTTATATTTCAATAGTTTCGAGAGAGAAATG--TGTATGTAACAG 262

DB 306 AACCAAGCTTACTTATCAACAAAGCTTGTGCGTGTGAGAGAAATGTTCTCAGTC 365

QY 263 AGGTGACCTTAAGAGTACTGACGCGTGTGCGCTCTCAGTACTTACATTAATTAAC 322

DB 366 ACGGTAAGCTAGGTCTTTTGTGACTCTGAGAACTTTTGAAGAAAGCTTAACAGT 425

QY 323 TTAGAACGTTGGTGTACTTTCAGTGAAGCTTACGTTGACTTTTATCGCGTATPAGC 382

DB 426 TCGGTTGTTTCTGCCGCACTTTTCAGAAAGACTATATCTTCGGAAGAAATACCGAG 485

QY 383 ACAATTCACCAACTCAAGCGCGCGGGAATTTGGGATTCAGAGTGAAGATTCGACT 442

DB 486 GAGACTGCTTCGATTTGTAAGACCAACCTGACGCTTACTCTGAGATCACTACT 545

QY 443 TAGCTGAGATTTTCTGGGTAATTGCCGAGCTCTCTGAATTAACAGCAAGTAGAAGA 502

DB 546 TAGCTGATGATTCATATCGACATCAACAGAACTTACTGACTACCAAAAGTGTCTGC 605

QY 503 TGTTCGGAGATATAGCGCTTAATAACTGA 533

DB 606 TGTGGCGCGGAAAGCGCACTCACAGAG 636

RESULT 14  
 LOCUS BYVMBPA 5980 bp RNA linear VRL 01-OCT-1993  
 DEFINITION Beet yellow virus (isolate German BYV-G) genes for putative  
 membrane-binding protein, heat shock 70-related protein, coat  
 protein homolog and coat protein.  
 ACCESSION X73475.1 GI:313689  
 VERSION X73475.1  
 KEYWORDS  
 ORGANISM  
 SOURCE  
 ORGANISM  
 Beet yellow virus  
 Beet yellow virus  
 Viruses; ssRNA positive-strand viruses, no DNA stage;  
 Closteroviridae; Closterovirus.

REFERENCE  
 AUTHORS Agranovsky, A.A., Koonin, E.Y., Boyko, V.P., Maiss, E., Lunina, N.A. and  
 Atabekov, J.G.  
 TITLE Beet yellow closterovirus: complete genome structure and  
 identification of a leader papain-like thiol protease  
 JOURNAL Unpublished  
 REFERENCE  
 AUTHORS Maiss, E.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-JUN-1993) E. Maiss, Inst. f. Biochemie und  
 Pflanzenvirologie, Biologische Bundesanstalt, Institut fuer  
 Biochemie, Messeweg 11/12 D-3300 Braunschweig, FRG

FEATURES  
 location/Qualifiers  
 source 1..5980  
 /organism="Beet yellow virus"  
 /mol\_type="genomic RNA"  
 /isolate="German (BYV0G)"  
 /db\_xref="taxon:12161"  
 /clone="pBYV-D22, pBYV-D31"  
 /note="unassigned protein product; putative membrane-binding  
 protein"

## CDS

/codon\_start=1  
 /protein\_id="CA51857.1"  
 /db\_xref="GI:313690"  
 /db\_xref="SPTREMBL:O65893"  
 /translation="YMDCLRSYLLAFGLICFLFCLVFIWFIYKQLFNNPESN"

CDS  
EARNSTTVV"  
235..2031  
/codon\_start=1  
/product="heat shock 70-related protein"  
/protein\_id="CA51858.1"  
/db\_xref="GI:313691"  
/db\_xref="GOA:Q65894"  
/db\_xref="SPTREMBL:Q65894"  
/translation="MYVPGIDPFGTTPSSVCAVYGEELVLFKORDSAVIPTFVFLHSDT  
QVAPRGVADAVLSDSSVGRGFRDLKRWIGDDEMYGDLKPKHYKTELLKXAOS  
SKSTKLDGSGVTPQATLPGILATFVYKALISTABAKCCCTGTCGVPANVACIO  
RSFTESCVMISGFCVYVWNEPSAALISCSRIKQATSVLYIDFGGTFDVSIAL  
NNTFVVASGDMGLGDRIDKAFVHLKKAQLPVNYKIDISFLKESLSKVSFLNF  
PVSEONVADVLYVNSVSELAEVAPFVEERTIKIVNEVYKRGSMLEPSVAKILNV  
GGSYLPGILSRLSVPEVECLVLPDAAVAGCALYACILRDSPLLVDCAAHN  
LSISYKCEIVCPVAGSPIPFTGRTVMNMGSAVYSAALFEDPFKCLNRIIF  
SGVVALGVDSVSGSNRTVPLTLEINVSAGITTFSLVGPVTKLVGNMAAYDSSY  
QLEERVADLHKNSDKVILTHALTYKPRKCTLDSDALFLKRLSADYRRKARKFS  
SYDDAVLNSSELLGRVLPKLRGSRVEKLDV"  
1937..3598  
/note="unnamed protein product; orf4"  
/codon\_start=1  
/protein\_id="CA51859.1"  
/db\_xref="GI:313692"  
/db\_xref="SPTREMBL:Q65895"  
/translation="WTRFSTPANYTWGELFRFPFGREKRLIMSEAAVRRPVYARD  
FRSNGVLSRKIFSDSTGESFVREFSLTLPKTYEVLGCVANEQALSNGNRISD  
YVNSLAVNVKTVGCKFNQVTEFEVKLNGNLAPSLVHGMSSLSNGELINPKD  
TKREFSLIPKQKDVIBSTDBAIVSSSLYLSHCLNLYETCNLSNSGKALYDEFLK  
YVIALYETSLDEHSPNDPLVAGVLYDMCFEYNTLSTYKLIIESFDGLSLYPLL  
SEIFSMWMBOPADPVLPELIDPAELLKPIINTLDTFLYKRLKRIEYESYEDSN  
ELIKRVDSLLTRDNPETLKAQRWVGHCTCYGVPRTACQKRYADAEVPCPLGERT  
INNSGVFEDELQKMPVSVRKRFQGSLSHAEISIFKRFVGVFPPIRLINPVYYS  
YLVNDYVHVYKRAQLQDELTLISNIEFDVAEMCCERVALQRRQOREKRPQGMKG  
VKNEVSPHARSIRVKKSNESLNIIMKDVGRQRLNPLFRKH"  
3546..4196  
/codon\_start=1  
/product="coat protein homolog"  
/protein\_id="CA51860.1"  
/db\_xref="GI:313693"  
/db\_xref="GOA:Q65896"  
/db\_xref="SPTREMBL:Q65896"  
/translation="MLASGRGLIHFTENTRDMETFFNSYDLAEYSEVNPKNLRK  
ETBELGIVIRERFSELYITDEDPVHLAFLIRANITTSVTNVVAGAEYVIGKK  
FLYVDAYVPELKECKMKPKMPVPCATFEDAYVIVARSLPKLPIKRTTGKGIIP  
SGIFLGADFLIATVCLNDHEKAIYVQASRAIDRAVSSVQKIVSLFDGLRSL"  
4267..4881  
/codon\_start=1  
/product="coat protein"  
/protein\_id="CA51861.1"  
/db\_xref="GI:313694"  
/db\_xref="GOA:Q03724"  
/db\_xref="SPTREMBL:Q03724"  
/translation="WGSABPISATVTEPNVSLADQTLHGEDCDKLRKNEBGLKIKG  
VPEBKGLALGLCLYSCATIGTSNKSVOPTSTFKSPFGGKBLTTHGELSLFDS  
QKLEGPKNKRCPCRFKDYISFAKEYGKRLPIRARHVLPMEDHVLADPFIIST  
STLELDQCKRLILARENATHTERSSESPTVSKQLORGLATGR"  
4878..5411  
/note="unnamed protein product; orf7"  
/codon\_start=1  
/protein\_id="CA51862.1"  
/db\_xref="GI:313695"  
/db\_xref="SPTREMBL:Q65897"  
/translation="MIGSVELAQTRPFRFVLLKGVFYVIAIETEESPEBAELPLVY  
LHDELVNKKGIKESYIDPKSCMRLLKSSVSATVSSGKSEDFSSSYSGKTES  
RIINRKVTFTENGTOLVFGMYGRDRCVSESEVLMFNVFVGAHCGTLPYCLNCELDK  
SGCELEITFSKNEVFL"  
5395..5928  
/note="unnamed protein product; orf8"  
/codon\_start=1  
/protein\_id="CA51863.1"  
/db\_xref="GI:313696"  
/db\_xref="SPTREMBL:Q65898"

ORIGIN  
/translation="MKFFENDGTSRAISRSESLLRKYZELGTNSPOSEVSEGINEFN  
ELARENLLVTVHEHREMMEXHPRQSSSELRPSLSEMLKXIRFLKRVVTPHAKETA  
SETLNAFLDEYCRJAGLTRDALREKWRKVSIVLFFHSHLELKEFVETENNFSFTLLK  
LNLISLRVSSQIIIGIAV"

Query Match	10.6%	Score 63	DB 14	Length 5980
Best Local Similarity	46.6%	Pred. No. 1.3e-08		
Matches 238	Conservative 0	Mismatches 270	Indels 3	Gaps 1
26	TTAGAACCTGGTATTAACGACGCTCTAGCTTAATGGTGTGACAGAAAGCTTTAT	85		
4304	TTGAAGAGTATGCTCGACAGCAACCGTGTTCACGGTGAAGACTGGACAACTAC	4263		
86	CTGCTGAAGTTGAAAATGTTGGTGCAGAAAGGCGCTCTTACAGAGGATTAAGATG	145		
4364	GGAAGAAATTTGAAAGAGTGTAAATTAAGAAAGGGGTTCGGAAGCAAACTCGTCTG	4423		
146	TGTTGGTCTACTCTCTTACGACACTCGCGGCAAGACACGCTCTTACAGTTCAGCGCG	205		
4424	CGTTAGACCTTTGTTGTTATCTCTGTCGACATAGTACTTTAATTAAGTTAGTCC	4483		
206	CAGATTCAAGCTTATATTTCAAAATGTTTGGAGAGAGAAATG--TGGTACTAACG	262		
4484	AACGACGCTACTTTCATCAAAAGCTTCTGCTGGTGGAGAGAAATGTTCTCTACTC	4543		
263	AGGTGACCTTAAGAGAGTACTCGACGGGTGTGGCCCTCGACATAGTTACTATAAAC	322		
4544	ACGGTGAAGTGAAGGCTTTCTTGAAGCTCTCGAAACCTTTGAAAGAAAGCTTAACAGT	4603		
323	TTAGAACCTGGTGTGCTACTTCACTGAGGCTTACGTTAGTCTTTGATCGCGTAAAGC	382		
4604	TCCGTTGTTTTCGCGGCACTTTTGAAGAGACTACATATCTTCGCGAAGAAATACCGAG	4663		
383	ACAAATTACCCCACTCAACCGCGGGGGAATGGGATTCGACGTAAATTCGACT	442		
4664	GAAAGCTGCTCCGATTCCTAGAGCAACGCTCACGGTCTACGCTGAAGATCACTACT	4723		
443	TAGCTGCAAGATTTTGTGGTACTTTCGCCGAAGCTCTCTGAATTCACGAAAGTGAAGA	502		
4724	TAGCTGCTGATTTCAATTCGACATCAAGAACTACTGACCTACCAAGAGTGTCTGCG	4783		
503	TGTTGCGAGATATGACCTCTTAATAACTGA	533		
4784	TGTTGGCGGCGGAAACCGCACTCACACAGA	4814		

RESULT 15  
AX655393/C  
LOCUS AX655393 2000 bp DNA linear PAT 22-MAR-2003  
DEFINITION Sequence 5263 from Patent WO03000898.  
ACCESSION AX655393  
VERSION AX655393.1 GI:29158207  
KEYWORDS  
SOURCE  
ORGANISM  
Oryza sativa  
Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzoideae; Oryza.

REFERENCE  
AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hsu, Y.M.,  
Katagiri, F., Quan, S., Tao, Y., Whitman, S., Xie, Z., Zhu, T. and Zou, G.  
TITLE Plant genes involved in defense against pathogens  
JOURNAL Patent: WO 03000898-A 5263 03-JAN-2003;  
SYNGENTA PARTICIPATIONS AG (CH)  
FEATURES  
source  
1..2000  
/organism="Oryza sativa"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4530"

ORIGIN  
Query Match 6.3%; Score 37.8; DB 6; Length 2000;

Best Local Similarity 9.4%; Pred. No. 1.7;  
Matches 48; Conservative 231; Mismatches 228; Indels 2; Gaps 1;

```
QY 16 GACAGAGAACTTGTGCACTGCTGTGATACCGACGCTTAGTCTTAATGCTGCGACAAG 75
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1009 SMSWTATMMKSKYKMTAISTWKMTAYKMRAYAMSRSRKTMCTGGRMATYCGTKM 950
QY 76 AAGCTTTTATCTGCTGAAGTTGAAAAATGTTGTGACAGAAAAGGGGCTCTTAACGAGGT 135
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
949 AAGRMWRBAMWCMYCCMMKMKWMTSCMMKMYRMTSCWYTMWGAARAYAYAMRRRRMTY 890
QY 136 ATGAAAGTGTGCTTCGCTTACTCTTACGCACTCGCGCAAGAACAGCTCTCTAG 195
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
889 KMSWRMYMTMTWAMTMTCMCAKMYMATGNATMMMYTMTYCYAMTCAKCKYKA 830
QY 196 GTTCAGCGCGCAGATTGACAGCTTAATATTTCAATAGTTTCGAGAGAGAAATGTGTA 255
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
829 MTKMWTWACAMRALISRMWMAAMGMRKRYKMRAYWWRWRCWKAGWALRMKSRYRMW 770
QY 256 GTAACAGAGGGTACCTTAAGAAAGTACTGACGGGTGTGCGCTCTCACTAGTCACT 315
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
769 KKYATRYMOMWMTMWSRMRKSYMWSGMRMRMSAMRYCSRMKCAKTKYASSARMT 710
QY 316 AATAAAGTGAAGCTTCGCTGCTTCACTGAGGCT--TACGTTGACTTTGTATCG 373
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
709 KRAQSYRFRRWYKWKGTTRIRYRSCMTTRABMSKRKRWAGASMSQWYMRGARS 650
QY 374 CGTATAAGCAAAATTACCCCACTCAAGCGCGCGGGAATTGGGATTCCAGTGAAG 433
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
649 MWYSKYSKSAKCKTKRMTSYMTSGMYGYSYKMSMTSKMSYVGRKMTCTMYTSMKG 590
QY 434 ATTGCTACTAGCTGCAGATTTTCTGGTACTTCCCGAAGCTCTGGAATTAAGGAAA 493
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
589 STRSRKMRSGMSRMTYRMRKRRKRYRMKCTWRMCYRWGYMTTTSRPM 530
QY 494 GTAGAAGATGTTGCGAGATATGTACGCT 522
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
529 YTGRYKARYTSKRRYMWYKRYKWCYYY 501
```

Search completed: April 7, 2004, 00:33:02  
Job time : 2193 secs



CC detection, and to transform host cells (especially *Agrobacterium vitis*,  
 CC *Agrobacterium tumefaciens*, grape, citrus, beet or tobacco cells) and  
 CC produce transgenic plants (claimed). They can be used to impart GRLAV-2  
 CC resistance to *Vitis* section or rootstock cultivars or *Nicotiana* cultivars  
 CC (claimed). Because extensive similarity exists between hsp70-related  
 CC sequence regions of GRLAV-2 and other closteroviruses, the DNA may also  
 CC be used to impart beet yellows virus resistance to beet cultivars or  
 CC *tristeza* virus resistance to citrus section cultivar/rootstock cultivars  
 CC (claimed). (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 597 BP; 165 A; 121 C; 157 G; 154 T; 0 U; 0 Other;

Query Match 100.0%; Score 597; DB 2; Length 597;  
 Best Local Similarity 100.0%; Pred. No. 6,4e-193;  
 Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAGATTGATGTCGACGACGACCTTACGCACTGTGTATACCGACGCTTACTCTA 60
DB 1 ATGAGATTGATGTCGACGACGACCTTACGCACTGTGTATACCGACGCTTACTCTA 60
QY 61 AATGATGTCGACGACGACCTTATCTGCTGATGTAAGAAAATGTTGTGCAAAAAGG 120
DB 61 AATGATGTCGACGACGACCTTATCTGCTGATGTAAGAAAATGTTGTGCAAAAAGG 120
QY 121 GCTCCTTAACGAGGGTATAGAAAGTGTGCTTACTCTTACGCACTGCGGCAAGA 180
DB 121 GCTCCTTAACGAGGGTATAGAAAGTGTGCTTACTCTTACGCACTGCGGCAAGA 180
QY 181 ACCACGCTCTTACGCTGACGCGGAGATTACAGCTTATATTTCATTAATGTTGCGA 240
DB 181 ACCACGCTCTTACGCTGACGCGGAGATTACAGCTTATATTTCATTAATGTTGCGA 240
QY 241 GAGAGGATGTGTAGTATACAGAGGTTGACCTTAAGAAAGTACTCGACGGTGTGCGCT 300
DB 241 GAGAGGATGTGTAGTATACAGAGGTTGACCTTAAGAAAGTACTCGACGGTGTGCGCT 300
QY 301 CTCACGTAGCTCATAAATTAAGAAAGTGTGCTGCTTACTGAGGAGCTTACTGCT 360
DB 301 CTCACGTAGCTCATAAATTAAGAAAGTGTGCTGCTTACTGAGGAGCTTACTGCT 360
QY 361 GACTTTTGTATCGGTATATACGAAATTAACCCCACTCAACGCGCGGCGGAATTGGGG 420
DB 361 GACTTTTGTATCGGTATATACGAAATTAACCCCACTCAACGCGCGGCGGAATTGGGG 420
QY 421 ATTCAGCTGAAGATTCGTACTTACCTGCAATTTTCTGGGTACTTGGCCGAAGCTCT 480
DB 421 ATTCAGCTGAAGATTCGTACTTACCTGCAATTTTCTGGGTACTTGGCCGAAGCTCT 480
QY 481 GAATTACAGCAAGTAGAAGATGTTCCGAGTATGTACGCTTAAATAATGAAGTGA 540
DB 481 GAATTACAGCAAGTAGAAGATGTTCCGAGTATGTACGCTTAAATAATGAAGTGA 540
QY 541 GTGTAAATATACCAAGTAGCAATTCGCTCAAGCTAGTAGAAGGAAGTATATTA 597
DB 541 GTGTAAATATACCAAGTAGCAATTCGCTCAAGCTAGTAGAAGGAAGTATATTA 597

```

## RESULT 2

AAV08874 standard; cDNA; 15500 BP.

AAV08874;

27-AUG-2003 (revised)  
 29-MAR-1999 (first entry)

XX Grapevine leafroll virus type 2 (GRLAV-2) genome.

XX GRLAV-2; closterovirus; grape; tobacco; transgenic plant;  
 XX disease resistance; virus resistance; beet yellows virus;  
 XX protease; methyltransferase; helicase; heat shock protein; coat protein;  
 XX RNA polymerase; se.

OS Grapevine leafroll virus.

XX	Key	Location/Qualifiers
XX	4.7923	
XX	/tag= a	"polypeptide (protease, methyltransferase, helicase)"
XX	/note= "Claim 26"	
XX	7922. .9301	
XX	/tag= b	"RNA-dependent RNA polymerase"
XX	/note= "Claim 29"	
XX	5365. .9535	
XX	/tag= c	"hydrophobic protein"
XX	/note= "Claim 43"	
XX	9551. .11350	
XX	/tag= d	"heat shock 70 protein"
XX	/note= "Claim 32"	
XX	11277. .12932	
XX	/tag= e	"heat shock 90 protein"
XX	/note= "Claim 35"	
XX	12844. .13515	
XX	/tag= f	"diverged coat protein"
XX	/note= "Claim 41"	
XX	13584. .14180	
XX	/tag= g	"coat protein"
XX	/note= "Claim 38"	
XX	14180. .14665	
XX	/tag= h	"undefined"
XX	/note= "Claim 45"	
XX	14667. .15284	
XX	/tag= i	"undefined"
XX	/note= "Claim 47"	
XX	15285. .15500	
XX	/tag= j	"Claim 48"
XX	/note= "Claim 48"	
XX	MO9853055-A1.	
XX	26-NOV-1998.	
XX	98MO-US010313.	
XX	20-MAY-1998;	
XX	97US-0047194P.	
XX	20-MAY-1997;	
XX	97US-0047194P.	
XX	(CORR ) CORNELL RES FOUND INC.	
XX	Zhu H. Ling K. Gonsalves D;	
XX	WPI, 1999-045307/04.	
XX	P-ESDB; AAW73476, AAW73477, AAW73478, AAW73479, AAW73480, AAW73481,	
XX	AAW73482, AAW73483, AAW73484.	
XX	Grapevine leafroll virus (type 2) proteins and polypeptides - and	
XX	encoding DNA, useful e.g. to impart grapevine leafroll resistance to	
XX	grape and tobacco plants and detect grapevine leafroll virus.	
XX	Disclosure; Page 12-20; 151pp; English.	
XX	A total of 15,500 bp of the RNA genome of grapevine leafroll virus type 2	
XX	(GRLAV-2) has been sequenced and cloned from GRLAV-2 isolated from	
XX	infected <i>Vitis vinifera</i> cv. Pinot Noir. About 85% of the total RNA genome	
XX	was revealed from 2 different clones. The sequence in the coat protein	
XX	gene region was determined and confirmed from several overlapping clones.	
XX	The genomic organization of GRLAV-2 includes 9 open reading frames (see	
XX	also AAV08864-72) encoding a polypeptide (see AAW73476) having papain-	



CC like protease, methyltransferase and helicase motifs, an RNA-dependent  
 CC RNA polymerase, heat shock proteins, coat proteins and other proteins of  
 CC unknown function (see AA07376-84). These proteins can be used to produce  
 CC antibodies, useful to detect GRV in samples e.g. by ELISA (claimed).  
 CC The nucleic acid molecules can be used to produce probes and primers for  
 CC such detection, and to transform host cells (especially Agrobacterium  
 CC vitis, Agrobacterium tumefaciens, grape, citrus, beet or tobacco cells).  
 CC and produce transgenic plants (claimed). In particular, they can be used  
 CC to impact GRV-2 resistance to Vitis scion or rootstock cultivars or  
 CC Nicotiana (claimed). Because extensive similarity exists between hsp70-  
 CC related sequence regions of GRV-2 and other closteroviruses, the DNA  
 CC may also be used to impart beet yellows virus resistance to beet  
 CC cultivars or tristeza virus resistance to citrus scion cultivar/rootstock  
 CC cultivars (claimed). (Updated on 27-AUG-2003 to correct OS field.)  
 CC  
 XX Sequence 15500 BP; 3984 A; 3178 C; 3970 G; 4368 T; 0 U; 0 Other;

Query Match 100.0%; Score 597; DB 2; Length 15500;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-192; Mismatches 0; Gaps 0;  
 Matches 597; Conservative 0; Indels 0;

DB 1 ATGAGATTGATGTCGACAGCAACCTTAGCACTGTGATTAACCGACCTCTAGTCTA 60  
 13584 ATGAGATTGATGTCGACAGCAACCTTAGCACTGTGATTAACCGACCTCTAGTCTA 13643  
 QY 61 AATGTTGTGACAAAGAGCTTTTATCTGCTGAAGTTGAAAAATGTTGTCAGAAAGGG 120  
 DB 13644 AATGTTGTGACAAAGAGCTTTTATCTGCTGAAGTTGAAAAATGTTGTCAGAAAGGG 13703  
 QY 121 GCTCCTAACGAGAGGTATAGAGTGTGCTCGCTCTCTTACGCACTCGCGGAGAA 180  
 DB 13704 GCTCCTAACGAGAGGTATAGAGTGTGCTCGCTCTCTTACGCACTCGCGGAGAA 13763  
 QY 181 ACCACGCTCTCTTAAAGTTTCAAGCGGAGATTCAGAGTTATATTTCAATATGTTCCGA 240  
 DB 13764 ACCACGCTCTCTTAAAGTTTCAAGCGGAGATTCAGAGTTATATTTCAATATGTTCCGA 13823  
 QY 241 GAGAGGAATGTGTATTAACAGAGGGTGACCTTAAGAGTTCTGACGGGTGTGGGCT 300  
 DB 13824 GAGAGGAATGTGTATTAACAGAGGGTGACCTTAAGAGTTCTGACGGGTGTGGGCT 13883  
 QY 301 CTCACCTAGGTTCACTATTAACCTTAGAAGCTTGGTGTGATCTTCACTGAGGCTTAAGCTT 360  
 DB 13884 CTCACCTAGGTTCACTATTAACCTTAGAAGCTTGGTGTGATCTTCACTGAGGCTTAAGCTT 13943  
 QY 361 GACTTTTGTATCCCGTATTAAGCAAAATTAACCCCACTCAAGCGCGCGGGAATTTGGGG 420  
 DB 13944 GACTTTTGTATCCCGTATTAAGCAAAATTAACCCCACTCAAGCGCGCGGGAATTTGGGG 14003  
 QY 421 ATTCGAGCTGAAGATTCGATTAAGTGTGAGATTTTCTGGGTAATTTGCCGGAAGCTCTCT 480  
 DB 14004 ATTCGAGCTGAAGATTCGATTAAGTGTGAGATTTTCTGGGTAATTTGCCGGAAGCTCTCT 14063  
 QY 481 GAATTTCAGGAAGTGAAGAGATTTTCCGAGTATGTAGCTTTAAAACTGAAGGTGGA 540  
 DB 14064 GAATTTCAGGAAGTGAAGAGATTTTCCGAGTATGTAGCTTTAAAACTGAAGGTGGA 14123  
 QY 541 GTGGTAAATACCAAGTGAAGCAATCTGCGTCACTAGTATGAAGGGAATTTATGTA 597  
 DB 14124 GTGGTAAATACCAAGTGAAGCAATCTGCGTCACTAGTATGAAGGGAATTTATGTA 14180

RESULT 3  
 AA087853  
 ID AA087853 standard; cDNA; 688 BP.

XX AA087853;  
 AC 16-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 02-NOV-1995 (first entry)  
 XX Sugar beet yellows virus capsid protein coding gene.

XX Sugar beet yellows virus capsid protein; transgenic plant; ss.  
 KW Beet yellows virus.  
 XX  
 OS  
 FH Key Location/Qualifiers  
 FT CDS 74..688  
 FT /\*tag= a  
 FT /product= "sugar beet yellows virus capsid protein"

RU2017820-Cl.

15-AUG-1994.

27-JUN-1991; 91SU-04950054.

27-JUN-1991; 91SU-04950054.

(IMMUNO) IMMUNOBIOTECH INST.

(BIOT) BIOYRCHN INST CO LTD.

Arganovskii AA, Boiko VP, Karasev AV;

WPI; 1995-113715/15.

P-PSDB; AAR72682.

Sugar beet yellows virus cDNA fragment encoding capsid protein - useful  
 for production of virus-resistant transgenic plants.

Claim 1; Col 7-10; 5pp; Russian.

The nucleotide sequence of the novel gene encoding the sugar beet yellows  
 virus (SBV) capsid protein. The gene encodes a protein of 204 amino  
 acids with mol. wt. 22.2 kD. The gene was obtained from reverse  
 transcribed RNA isolated from purified SBV. The fragment is useful for  
 the production of virus resistant transgenic plants by genetic  
 engineering methods. (Updated on 25-MAR-2003 to correct PN field.)  
 CC (Updated on 25-MAR-2003 to correct PF field.) (Updated on 16-OCT-2003 to  
 standardise OS field)

Sequence 688 BP; 190 A; 155 C; 157 G; 186 T; 0 U; 0 Other;

Query Match 10.7%; Score 53.8; DB 2; Length 688;  
 Best Local Similarity 48.1%; Pred. No. 7.8e-11; Mismatches 227; Indels 3; Gaps 1;  
 Matches 213; Conservative 0;

QY 95 TTGAAAAAATGTGTGTCAGAAAAGGGCTCTTAACGAGGATATAGATGTGTGCTC 154  
 DB 180 TCGAAGAGTGTGTAATTAAGAGGGTTCGGAAGATTAACCTCGAATGCGTTAGGAC 239  
 QY 155 TACTCTTTAGCACTCGCGCAAGAACCAAGTCTCTTAAGTTCA---GCGGCGAGATT 211  
 DB 240 TTTGTTTGTATTCCTGTGCTACGATAGGCACTTCCAAAGTTAAAGTCCCAACCGAGT 299  
 QY 212 CAGACGTTATATTTTCAATATAGTTTCGAGAGAGAAATGTGTATTAACAGAGGTGAC 271  
 DB 300 CTACCTTCATCAAGCTTGT 359  
 QY 272 TTAAGAAAGTACTCGAGCGGCTGCGCTCTCACTAGGTTCACTAATAAATTGAAGCT 331  
 DB 360 TGAATTCCTTCTGTGGGTCTCAAAAATTTTGGAGGGAATAACCTAAATTTGGGGTGT 419  
 QY 332 TCGGTCTACTTTTCACTGAGGCTTAAGTTATTTTGTATGCGGTATTAAGCAAAATTAC 391  
 DB 420 TCTGCGTACTTTTCAAGAGGACTACATATTCCTTCCCAAGGAATTAACGAGGGAATTAAC 479  
 QY 392 CCGAAGTCAAGCGCGCGGGAATTTGGGGAATTCAGGCTGAAGATTCTACTTACTGCTCAG 451  
 DB 480 CTCGATTTGCGAGAGCTTAACCGTACGCTCTACCGGTGAAGTACATTAATGACCGCTG 539  
 QY 452 ATTTTCTGGGTACTTGTCCCGAAGCTCTGTGAATTAACGAAAGTGAAGATGTTCCGGA 511  
 DB 540 ACTTCATATGACGCTGACGGAAGTCACTGACCTGACCTGACCTGACCTGACCTGACCTGAC 599

QY 512 GTATGTACGCTCTTAAATCTGAA 534  
 Db 600 GCGAAAAGCCACTCAGCGGAA 622

## RESULT 4

AAAT42975 standard; cDNA; 783 BP.

AAAT42975;

16-OCT-2003 (revised)

14-JAN-1997 (first entry)

BYV coat protein cDNA.

XX Tobacco mosaic virus; coat protein; beet yellow virus; BYV;  
 KW beet western yellow virus; BWYV; beet necrotic yellow vein virus; BNYYV;  
 KW luteovirus; closterovirus; sugar beet; fruit; mango; apple; pear; banana;  
 KW field crop; sunflower; wheat; barley; maize; vegetable; potato; carrot;  
 KW cabbage; onion; ss.

XX Beet yellows virus.

XX Key Location/Qualifiers

FT sig\_peptide 13..78

FT /tag= a /note= "Omega sequence from TMV"

FT 89..703

FT /tag= b /product= "BYV\_coat\_protein"

XX WO9428147-A1.

XX 08-DEC-1994.

XX 01-JUN-1994; 94WO-EP001786.

XX 02-JUN-1993; 93GB-00011332.

XX (SANO ) SANDOZ LTD.

XX (SANO ) SANDOZ PATENT GMBH.

XX (SANO ) SANDOZ-ERFINDUNGEN VERM GRS MBH.

XX Bojsen K, Brunstedt J;

XX WPI, 1995-022818/03.

XX Recombinant DNA sequence comprising at least two coat protein genes -  
 used to combat viruses in e.g. sugar beet.

XX Disclosure: Page 11-12; 23pp; English.

XX The sequences given in AA142975-77 represent cDNA sequences encoding coat  
 CC proteins from beet yellow virus (BYV), beet western yellow virus, (BWYV)  
 CC and beet necrotic yellow vein virus (BNYYV) respectively. These sequences  
 CC are used in the recombinant DNA sequence of the invention. This comprises  
 CC nucleotide sequences encoding the coat proteins of at least one  
 CC luteovirus and at least one closterovirus. This recombinant DNA may be  
 CC used to transform sugar beet for combating viral infection. Other plants  
 CC may also be transformed e.g. fruit such as mangoes, apples, pears, and  
 CC bananas, and field crops such as sunflower, wheat, barley, maize, and  
 CC vegetables such as potatoes, carrots, cabbage and onion. (updated on 16-  
 CC OCT-2003 to standardise OS field)

XX Sequence 783 BP; 223 A; 178 C; 171 G; 211 T; 0 U; 0 Other;

XX Query Match 10.6%; Score 63; DB 2; Length 783;

XX Best Local Similarity 46.6%; Pred.No. 1.6e-10;  
 Matches 238; Conservative 0; Mismatches 270; Indels 3; Gaps 1;

QY 26 TTAGCAACTGGTGAATACCGAGCGCTTAGTCTAATGATGTGCACAAAGAGCTTTAT 85

Db 126 TTGAAAACGTGAGTCTCGCAGACCAACGTGTTTGACGCGTAGAAGACTGCGACAAACTAC 185  
 QY 86 CTGCTGAAGTTGAAAATAATGTTGTCGAGAAAGGGGCTCTTAAGAGGGTATAGAGTGG 145  
 Db 186 GGAGAAATTTGGAAGAGTGTGAAATTTGAAATTTGAAAGGGGTTCCGAAAGACAAACTCGTCTCG 245  
 QY 146 TGTTCGCTACTCTCTTTACGCACTCGCGGAGAACCAACGCTCTTAAGGTTCAAGCGCG 205  
 Db 246 CGTTGAACTTGTGTTGTAATTCCTGTGACAGCATAGTACTTCTTAATTAAGTTAGTGTC 305  
 QY 206 CAGATTCAAGCGTTATATTTTCAATTAAGTTTCGAGAGAGGAAG--TGGTACTAACG 262  
 Db 306 AACCGACGTCTACTTTCATCAAAAGCTTGCTGCGTGGAGAGAAATGTTCTCTACTC 365  
 QY 263 AAGGTGACCTTAAGAAAGGACTCGACGGGAGTGGCGCTCTCACTAGTTCACTAATAAC 322  
 Db 366 ACGGTGAAGTGAAGTCTTTCTGACCTCTCAAAACCTTTGAAAGGAAGCCTAACAAAGT 425  
 QY 323 TTAGAACGTTGCGTGTGTACTTCACTGAGGCTTACGTTGATTCGCGTATAGC 382  
 Db 426 TCGGTGTTTCTGCCGCACTTTTCAGAAAGACTACATATCTTTCGCGAAGGAATACCGAG 485  
 QY 383 ACAATTTACCCCACTCAACAGCGCGGGAATTGGGATTCGAGCTGAAGATTGCTACT 442  
 Db 486 GAAAGCTGCTCTCGATTGCTGAGGCAACCGTCAAGGCTTACTGTAAGATCACTACT 545  
 QY 443 TAGCTGCAATTTTCTGGTACTTCCCGAAGCTCTTGAAATTACAGCAAAAGTAGAAGA 502  
 Db 546 TAGCTGCTGATTTCAATGACATCAAGAACTTACGACTACACAAAGTGTGCTGCG 605  
 QY 503 TGTTCGCGAGTATGATAGCTCTAATAACTGA 533  
 Db 606 TGTTCGCGCGCAAAAGCCACTCAGCAGCA 636

## RESULT 5

ADA71938/c

ID ADA71938 standard; DNA; 2000 BP.

XX ADA71938;

XX 20-NOV-2003 (first entry)

XX Rice gene, SEQ ID 5263.

XX Plant; bacterial infection; fungal infection; viral infection; rice;

XX gene; ds.

XX Oryza sativa.

XX WO2003000898-A1.

XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX (SYGN ) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

XX Katsagiri F, Qian S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to  
 PT pathogenic infection for conferring resistance or tolerance to a plant to  
 PT bacterial, fungal or viral infection by determining or detecting plant  
 PT gene expression.

XX Claim 27; SEQ ID NO 5263; 899pp; English.

CC The present invention relates to a method (M1) for identifying genes  
 CC involved in plant resistance or response to pathogenic infection. M1  
 CC comprises identifying a gene whose expression is significantly altered in  
 CC the incompatible interaction of plant gene expression relative to  
 CC the expression of the gene in an uninfected plant, in a mutant plant that  
 CC does not express a gene associated with response to pathogenic infection,  
 CC or in a corresponding incompatible or compatible interaction. (M1) is  
 CC useful for conferring resistance to resistance or tolerance to a plant to  
 CC bacterial, fungal or viral infection. The present sequence was used to  
 CC illustrate the invention.

XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

XX Query Match 6.3%; Score 37.8; DB 7; Length 2000;

XX Best Local Similarity 9.4%; Pred. No. 0.1; Mismatches 228; Indels 2; Gaps 1;

XX Matches 48; Conservative 231; Mismatches 228; Indels 2; Gaps 1;

QY 16 GACAGCAACCTTAGCACTGGATGATACCGACGCTTAGTCTTAATGGTGTGACAG 75

DB 1009 SMSWTYAMWKYTKYMTAYSSTWKMYAYMSRSRKTWCTGKEMATYCGTKW 950

QY 76 AAGCTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 135

DB 949 AAGRMWRBAAKMYCCMKMKKTSCKMKMKYMTWGMARVYAMRRRTY 890

QY 136 ATGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 195

DB 889 KWSRMYMTWTKWATWMTWCMAMVATGATWMTWMTWMTWMTWMTWMTWMTWMT 830

QY 196 GTTCAGCGGCGGATTCAGCGCTATATTTTAAATGATTTGGAGAGGAATGGTA 255

DB 829 MKKMTWTTACAMRATSRWRBAAAGMKKRYKMRATWMTWMTWMTWMTWMTWMTWMT 770

QY 256 GTAACAGAGGCTGACCTTAGAAGTACTGACGGGTGGCTCTCTCACTAGTTCA 315

DB 769 KXATATYKMMAMTMMWMSRWKSYRMSGMRMSAMRYCSRMKCAKTKYASAR 710

QY 316 AATAAATGAAAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 373

DB 709 KKAARSRYRRRMYRKRMTATYRWRSCMTFRAMSRKRKRAAGSMKSMWYRGA 650

QY 374 CGTATAGCAACAATTAACCACTCAACGCGCGCGGGAATGGGATTCAGCTGA 433

DB 649 MMYSKXSCAKCKKTRMTSSYMTGMYSSYSMSMTSKMSMGMTCTMTWTSKG 590

QY 434 ATTCGTAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 493

DB 589 STSRSDGMRSGMSRMYRMRKMRKRYRMYRMYRMYRMYRMYRMYRMYRMYRMY 530

QY 494 GTAGAAAGATGTCGCGAGATGATGCT 522

DB 529 YTGRIYKARITSKRMYMYKRYKWCWYYY 501

RESULT 6  
 ABN67902 standard; DNA; 1332 BP.

XX ABN67902;

XX 01-UTL-2002 (first entry)

XX Streptococcus polynucleotide SEQ ID NO 3717.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

XX group A streptococcus; Streptococcus pyogenes; antibacterial; gene;

XX anti-inflammatory; infection; vaccine; meningitis; gene therapy; ds.

XX Streptococcus agalactiae.

XX MO200234771-A2.

XX 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB004789.

XX 27-OCT-2000; 2000GB-00026333.

XX 24-NOV-2000; 2000GB-00028727.

XX 07-MAR-2001; 2001GB-00005640.

XX (CHIR-) CHIRON SPA.

XX (GENO-) INST GENOMIC RES.

XX Telford U, Massignani V, Margarit Y Rostl, Grandi G, Frazer C;

XX Tectelin H;

XX WPI; 2002-352536/38.

XX P-PSDB; ABP27271.

XX New Streptococcus protein for the treatment or prevention of infection or

XX disease caused by Streptococcus bacteria, such as meningitis, and for

XX detecting a compound that binds to the protein.

XX Claim 7; Page 3526; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B

XX streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS

XX (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

XX the specification. The proteins have antibacterial and anti-inflammatory

XX activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and

XX antibodies that bind (I) are used in the manufacture of medicaments for

XX the treatment or prevention of infection or disease caused by

XX Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

XX Nucleic acids encoding (I) are used to detect Streptococcus in a

XX biological sample. (I) is used to determine whether a compound binds to

XX (I). A composition comprising (I) or a nucleic acid encoding (I), may be

XX used as a vaccine or diagnostic composition. The disease caused by

XX Streptococcus that is prevented or treated may be meningitis. Nucleic

XX acid encoding (I) may be used to recombinantly produce (I) and may be

XX used in gene therapy. Antibodies to (I) are used for affinity

XX chromatography, immunoassays, and distinguishing/identifying

XX Streptococcus proteins

XX Sequence 1332 BP; 423 A; 205 C; 278 G; 426 T; 0 U; 0 Other;

XX Query Match 6.2%; Score 36.8; DB 6; Length 1332;

XX Best Local Similarity 51.9%; Pred. No. 0.18; Mismatches 97; Indels 3; Gaps 1;

XX Matches 108; Conservative 0; Mismatches 97; Indels 3; Gaps 1;

QY 55 AGCTTAATGCTGTCGACAAAGACCTTTATCTGCTGAAGTTGAAAAAATGTTGTCAG 114

DB 517 ACTTTGATGATGCTGACGAAAGATGACACGTGATGAAATGAAATGATGACGAAT 576

QY 115 AAAGGGGCTCTTACAGAGGATATGAGTGTGCTGCTACTCTTACGCACTGCG 174

DB 577 AGTGAAGAACTTTGGAAGCTGGAAGAAATGAGATGTCGAAGGATTTCTGCTGAT 636

QY 175 GCAAGAACCACTGCTC--TAAGTTACGCGGCGAGATTGAGAGTTATTTTCAAT 231

DB 637 GAAATGATGGCGCTGAAGTATGTTGCTCCGCGCACTGATCTTCATGATGATCAAC 696

QY 232 AGTTGCGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 259

DB 697 AATGATGACATCAATTAATTGAAAGAA 724

RESULT 7

ABN71527\_13/c

Continuation (14 of 22) of ABN71527 from base 130001 (Streptococcus polynucleotide SEQ 1)

WP Sequence split into 22 fragments LOCUS ABN71527 Accession ABN71527

WP Fragment Name Begin End

WP ABN71527\_00 1 110000

WP ABN71527\_01 100001 210000

WP ABN71527\_02 200001 310000

WP ABN71527\_03 300001 410000

WP ABN71527\_04 400001 510000

WP ABN71527\_05 500001 610000  
 WP ABN71527\_06 600001 710000  
 WP ABN71527\_07 700001 810000  
 WP ABN71527\_08 800001 910000  
 WP ABN71527\_09 900001 1010000  
 WP ABN71527\_10 1000001 1110000  
 WP ABN71527\_11 1100001 1210000  
 WP ABN71527\_12 1200001 1310000  
 WP ABN71527\_13 1300001 1410000  
 WP ABN71527\_14 1400001 1510000  
 WP ABN71527\_15 1500001 1610000  
 WP ABN71527\_16 1600001 1710000  
 WP ABN71527\_17 1700001 1810000  
 WP ABN71527\_18 1800001 1910000  
 WP ABN71527\_19 1900001 2010000  
 WP ABN71527\_20 2000001 2110000  
 WP ABN71527\_21 2100001 2155561

Query Match 6.2%; Score 36.8; DB 6; Length 110000;  
 Best Local Similarity 51.9%; Pred. No. 1.7; Indels 3; Gaps 1;  
 Matches 108; Conservative 0; Mismatches 97;

QY 55 AGTCTAATGATGTCGACAGAGCTTTATCTGCTGAAGTTGAAAAAGTTGGTGCAG 114  
 Db 76027 ACTTTGATGATGCTGACGAGAGATGACACGCTGATGAAATTTGATTTGACGAAAT 75968  
 QY 115 AAAAGGGCTCTTAACGAGGGTATGAAAGTGTGCTTACTCTTTACGCACTGCG 174  
 Db 75967 AGTGAAGAACTTTGAAAGCTGAAAGATGCTGCAAGGATTTCTCGCTAGAT 75908  
 QY 175 GCAGAGAACACAGTCTCC---TAAGTTACAGCGGCGAGATTCAGCGTTATTTCAAT 231  
 Db 75907 GAAATGATGCGCGCTGAAGTTATGTTTCGCGCACTGATCTTCTGATGATCAATCAAC 75848  
 QY 232 AGTTTCGAGAGAGGAATGTGTAGTAA 259  
 Db 75847 AATGATGCAATCAATATTTGAAGAA 75820

RESULT 8  
 ACA38387/c  
 ID ACA38387 standard; DNA; 1605 BP.  
 XX ACA38387;  
 AC ACA38387;  
 XX 19-JUN-2003 (first entry)  
 DT 19-JUN-2003 (first entry)  
 XX Prokaryotic essential gene #20044.  
 DE Prokaryotic essential gene; cell proliferation;  
 XX Antisense; ds; prokaryotic essential gene; cell proliferation;  
 KM drug design; gene.  
 XX Mycobacterium bovis.  
 OS Mycobacterium bovis.  
 PN W0200277183-A2.  
 XX W0200277183-A2.  
 PD 03-OCT-2002.  
 XX 03-OCT-2002.  
 PF 21-MAR-2002; 2002WO-US009107.  
 XX 21-MAR-2002; 2002WO-US009107.  
 PR 21-MAR-2001; 2001US-00915242.  
 PR 06-SEP-2001; 2001US-00948933.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX 06-MAR-2002; 2002US-0362699P.  
 PA (ELIT-) ELITRA PHARM INC.  
 XX (ELIT-) ELITRA PHARM INC.  
 PI Marg L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW,  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth KA, Xu HH;  
 DR WPI; 2003-029926/02.  
 DR P-PSDB; ABU34517.

XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.

Claim 14, SEQ ID NO 26257, 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway of the gene product or that inhibits cellular proliferation; (8) required for proliferation, or that inhibits cellular proliferation; (9) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-regulated gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

SO Sequence 1605 BP; 334 A; 490 C; 516 G; 265 T; 0 U; 0 Other;

Query Match 6.1%; Score 36.6; DB 7; Length 1605;  
 Best Local Similarity 57.4%; Pred. No. 0.23; Indels 0; Gaps 0;  
 Matches 66; Conservative 0; Mismatches 49;

QY 184 ACGTCTCCTAAGTTCAGCGCGGAGATTGACGTTATTTCAATTAATTTGAGAG 243  
 Db 200 ACGGTATTTGCGCGGAGATGACCGTGGACTCGCATTTGTTGGTTCTCGGACAG 141  
 QY 244 AGAATGTGTAGTAAACAGAGGTTGACCTTAAGAAAGTACTGACGGGTGTGCC 298  
 Db 140 GGTGGGTGCTAGGAGACAGATTGTGTTACTGACGAGGCTACTGACGCGGTACACGC 86

RESULT 9  
 AA199682\_18/c  
 Continuation (19 of 45) of AA199682 from base 1800001 (Mycobacterium tuberculosis strain  
 WP Sequence split into 45 fragments LOCUS AA199682 Accession AA199682  
 WP Fragment Name Begin End  
 WP AA199682\_00 1 110000  
 WP AA199682\_01 100001 210000  
 WP AA199682\_02 200001 310000  
 WP AA199682\_03 300001 410000  
 WP AA199682\_04 400001 510000  
 WP AA199682\_05 500001 610000  
 WP AA199682\_06 600001 710000  
 WP AA199682\_07 700001 810000  
 WP AA199682\_08 800001 910000  
 WP AA199682\_09 900001 1010000  
 WP AA199682\_10 1000001 1110000  
 WP AA199682\_11 1100001 1210000  
 WP AA199682\_12 1200001 1310000  
 WP AA199682\_13 1300001 1410000

```

WP AA199682_14 1400001 1510000
WP AA199682_15 1500001 1610000
WP AA199682_16 1600001 1710000
WP AA199682_17 1700001 1810000
WP AA199682_18 1800001 1910000
WP AA199682_19 1900001 2010000
WP AA199682_20 2000001 2110000
WP AA199682_21 2100001 2210000
WP AA199682_22 2200001 2310000
WP AA199682_23 2300001 2410000
WP AA199682_24 2400001 2510000
WP AA199682_25 2500001 2610000
WP AA199682_26 2600001 2710000
WP AA199682_27 2700001 2810000
WP AA199682_28 2800001 2910000
WP AA199682_29 2900001 3010000
WP AA199682_30 3000001 3110000
WP AA199682_31 3100001 3210000
WP AA199682_32 3200001 3310000
WP AA199682_33 3300001 3410000
WP AA199682_34 3400001 3510000
WP AA199682_35 3500001 3610000
WP AA199682_36 3600001 3710000
WP AA199682_37 3700001 3810000
WP AA199682_38 3800001 3910000
WP AA199682_39 3900001 4010000
WP AA199682_40 4000001 4110000
WP AA199682_41 4100001 4210000
WP AA199682_42 4200001 4310000
WP AA199682_43 4300001 4410000
WP AA199682_44 4400001 4411529

```

Query Match Best Local Similarity 57.4%; Score 36.6; DB 4; Length 110000;  
Matches 66; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

```

Qy 184 ACGTCTCTTAAGTTCACGCGCAGATTACAGCTTATTTCAATAAGTTTCGAGAG 243
Db 33577 ACGGCTACTTCCGCGCGAGGTACGCGTGGGACTCGCATATTGTTGGTCTCCGAGACG 33518
Qy 244 AGGAATGTGTAGTACAGAGGGTGACCTTAGAAGTACTCGACGGGTGTGGC 298
Db 33517 GGTGGGTGTAGGAGACAGATTGTGACTGACGAGCTACTCGACGGGTACACGC 33463

```

RESULT 10  
AA199683\_18/c  
Continuation (19 of 44) of AA199683 from base 1800001 (Mycobacterium tuberculosis strain  
WP Sequence split into 44 fragments Locus AA199683 Accession AA199683

Fragment Name	Begin	End
WP AA199683_00	1	110000
WP AA199683_01	100001	210000
WP AA199683_02	200001	310000
WP AA199683_03	300001	410000
WP AA199683_04	400001	510000
WP AA199683_05	500001	610000
WP AA199683_06	600001	710000
WP AA199683_07	700001	810000
WP AA199683_08	800001	910000
WP AA199683_09	900001	1010000
WP AA199683_10	1000001	1110000
WP AA199683_11	1100001	1210000
WP AA199683_12	1200001	1310000
WP AA199683_13	1300001	1410000
WP AA199683_14	1400001	1510000
WP AA199683_15	1500001	1610000
WP AA199683_16	1600001	1710000
WP AA199683_17	1700001	1810000
WP AA199683_18	1800001	1910000
WP AA199683_19	1900001	2010000
WP AA199683_20	2000001	2110000
WP AA199683_21	2100001	2210000
WP AA199683_22	2200001	2310000

```

WP AA199683_23 2300001 2410000
WP AA199683_24 2400001 2510000
WP AA199683_25 2500001 2610000
WP AA199683_26 2600001 2710000
WP AA199683_27 2700001 2810000
WP AA199683_28 2800001 2910000
WP AA199683_29 2900001 3010000
WP AA199683_30 3000001 3110000
WP AA199683_31 3100001 3210000
WP AA199683_32 3200001 3310000
WP AA199683_33 3300001 3410000
WP AA199683_34 3400001 3510000
WP AA199683_35 3500001 3610000
WP AA199683_36 3600001 3710000
WP AA199683_37 3700001 3810000
WP AA199683_38 3800001 3910000
WP AA199683_39 3900001 4010000
WP AA199683_40 4000001 4110000
WP AA199683_41 4100001 4210000
WP AA199683_42 4200001 4310000
WP AA199683_43 4300001 4403765

```

Query Match Best Local Similarity 57.4%; Score 36.6; DB 4; Length 110000;  
Matches 66; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

```

Qy 184 ACGTCTCTTAAGTTCACGCGCAGATTACAGCTTATTTCAATAAGTTTCGAGAG 243
Db 24455 ACGGCTACTTCCGCGCGAGGTACGCGTGGGACTCGCATATTGTTGGTCTCCGAGACG 24396
Qy 244 AGGAATGTGTAGTACAGAGGGTGACCTTAGAAGTACTCGACGGGTGTGGC 298
Db 24395 GGTGGGTGTAGGAGACAGATTGTGACTGACGAGCTACTCGACGGGTACACGC 24341

```

RESULT 11  
ABX74308  
ID ABX74308 standard; DNA; 593 BP.  
XX  
AC ABX74308;  
XX  
DT 18-MAR-2003 (first entry)  
XX  
DE DNA sequence #80 encoding M. tuberculosis antigenic polypeptide.  
XX  
KM Mycobacterium tuberculosis antigenic polypeptide; immune response;  
XX tuberculosis infection; gene therapy; gene; ds.  
XX  
OS Mycobacterium tuberculosis.  
XX  
PN US6465633-B1.  
XX  
PD 15-OCT-2002.  
XX  
PF 23-DEC-1999; 99US-00470191.  
XX  
PR 24-DEC-1998; 98US-0113952P.  
XX  
PA (COR-) CORIXA CORP.  
XX  
PI Skelky Y;  
XX  
DR WPI, 2003-147072/14.  
XX  
XX  
XX Novel isolated mycobacterial polynucleotide, useful for treating,  
XX preventing or diagnosing Mycobacterium tuberculosis infection, for  
XX producing Mycobacterium tuberculosis secretory polypeptides and DNA  
XX vaccines.  
XX  
PS Example; Col 81-82; 48bp; English.  
XX  
XX The present invention relates to the isolation of polynucleotide  
XX sequences encoding Mycobacterium tuberculosis antigenic polypeptides. The

CC polynucleotide sequences of the invention are useful for treating,  
CC preventing, and diagnosing M. tuberculosis infection, for producing M.  
CC tuberculosis secretory polypeptides, for producing DNA vaccines, for  
CC diagnostic purposes, as molecular probes or primers to detect the  
CC presence of bacteria in a biological sample, for inducing and/or  
CC enhancing immune responses to M. tuberculosis, and in gene therapy.  
CC ABX74229-ABX74319 represent DNA sequences encoding M. tuberculosis  
CC antigenic polypeptides

XX Sequence 593 BP; 95 A; 181 C; 186 G; 120 T; 0 U; 11 Other;  
SQ

Query Match 6.0%; Score 35.6; DB 7; Length 593;  
Best Local Similarity 56.5%; Pred. No. 0.31;  
Matches 65; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 184 ACGGCTCTAGAGTTCAGCGCGAGATTACAGCTTATTTCAATAGTTCCGAGAG 243  
Db 101 ACGGCTCTAGTTCGCGAGGAGTACCGTGGACTCGCATTTGTTGGTTCTCCGACAG 160

QY 244 AGGATGTGTAGTACAGAGAGGTGACCTTAGAGGTACTCGAGGTTGCGC 298  
Db 161 GGTGGTGTGATGACGAGTGTGACTTACCTACGAGGCTACTCGAGGTTACAGGC 215

RESULT 12  
AAL04577  
ID AAL04577 standard; DNA; 6756 BP.

XX AAL04577;  
XX  
XX 21-NOV-2001 (first entry)  
XX  
XX Human reproductive system related antigen DNA SEQ ID NO: 7265.  
XX Human reproductive system related antigen; reproductive system disorder;  
XX cancer; gene therapy; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200155320-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001339.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
XX 04-FEB-2000; 2000US-0180628P.  
XX 24-FEB-2000; 2000US-0184664P.  
XX 02-MAR-2000; 2000US-0186350P.  
XX 16-MAR-2000; 2000US-0189874P.  
XX 17-MAR-2000; 2000US-0190076P.  
XX 18-APR-2000; 2000US-0198123P.  
XX 19-MAY-2000; 2000US-0205515P.  
XX 07-JUN-2000; 2000US-0209467P.  
XX 28-JUN-2000; 2000US-0214886P.  
XX 30-JUN-2000; 2000US-0215135P.  
XX 07-JUL-2000; 2000US-0216647P.  
XX 07-JUL-2000; 2000US-0216880P.  
XX 11-JUL-2000; 2000US-0217487P.  
XX 11-JUL-2000; 2000US-0217496P.  
XX 14-JUL-2000; 2000US-0218290P.  
XX 26-JUL-2000; 2000US-0220963P.  
XX 26-JUL-2000; 2000US-0220964P.  
XX 14-AUG-2000; 2000US-0224518P.  
XX 14-AUG-2000; 2000US-0224519P.  
XX 14-AUG-2000; 2000US-0225213P.  
XX 14-AUG-2000; 2000US-0225214P.  
XX 14-AUG-2000; 2000US-0225267P.  
XX 14-AUG-2000; 2000US-0225267P.  
XX 14-AUG-2000; 2000US-0225268P.  
XX 14-AUG-2000; 2000US-0225270P.  
XX 14-AUG-2000; 2000US-0225447P.  
XX 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229349P.  
PR 01-SEP-2000; 2000US-0229349P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235844P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 29-SEP-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0244674P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.

PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	08-NOV-2000;	2000US-0246613P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249246P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0256719P.
PR	08-DEC-2000;	2000US-0251866P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259678P.
XX		
PA	(HUMA-)	HUMAN GENOME SCL INC.
XX		
PI	Rosen CA,	Barash SC, Ruben SM;
DR	WPI;	2001-465570/50.
XX		
PT	Isolated nucleic acid molecule encoding a reproductive system antigen is	
PT	used in preventing, treating or ameliorating a medical condition.	
XX		
PS	Disclosure; SEQ ID NO 7265; 1297bp + Sequence Listing; English.	
XX		
CC	The present invention provides the protein and coding sequences of a	
CC	number of human reproductive system related antigens. These can be used	
CC	in the prevention and treatment of reproductive system disorders,	
CC	including cancer. The present sequence is a genomic sequence encoding a	
CC	protein of the invention	
XX		
SQ	Sequence 6756 BP; 1603 A; 1601 C; 1877 G; 1675 T; 0 U; 0 Other;	
XX		
QY	Query Match	5.7%; Score 33.8; DB 4; Length 6756;
Db	Best Local Similarity	56.9%; Pred. No. 4.3; Mismatches 0; Gaps 0;
QY	Matches 62; Conservative 0; Mismatches 47; Indels 0; Gaps 0;	
Db	218 TTATATTTTCAATAGTTCGGAGAGAGGAATGGTAGTAACAGAGGTCACCTTAGA 277	
	1737 TTATGATTTTTCATTATTTAGACAGAAAGAAATGTAGTGGAGAGCAGGTCATGATTAAG 1796	
QY	278 AGGACTGAGCGGGTGGCGGCTCTCAGCTAGGTCACAAATAACTTAG 326	
Db	1797 AGAGGTGCGAAGACATGTGAGCAAAAGAAATTTATGTCTATTAATTAG 1845	
RESULT 13		
ABL97500		
ABL97500	standard; DNA; 6756 BP.	
ABL97500;		

XX	DT	21-JUN-2002	(first entry)
XX	DE	Human testicular antigen encoding DNA fragment SEQ ID NO: 2152.	
XX	XX	Human, testicular antigen; testes; cancer; metastasis; immune disorder;	
XX	KW	reproductive system disorder; urinary system disorder; gene therapy;	
XX	KW	cardiovascular disorder; respiratory disorder; neurological disorder;	
XX	KW	gastrointestinal disease; infection; cytostatic; gene; ds.	
XX	OS	Homo sapiens.	
XX	XX	WO200155317-A2.	
XX	PD	02-AUG-2001.	
XX	PF	17-JAN-2001, 2001WO-US001329.	
XX	PR	31-JAN-2000, 2000US-0179065P.	
XX	PR	04-FEB-2000, 2000US-0180628P.	
XX	PR	24-FEB-2000, 2000US-0184644P.	
XX	PR	02-MAR-2000, 2000US-0186350P.	
XX	PR	15-MAR-2000, 2000US-0188874P.	
XX	PR	17-MAR-2000, 2000US-0190076P.	
XX	PR	18-APR-2000, 2000US-0198123P.	
XX	PR	19-MAY-2000, 2000US-0205151P.	
XX	PR	07-JUN-2000, 2000US-0209467P.	
XX	PR	28-JUN-2000, 2000US-0214886P.	
XX	PR	30-JUN-2000, 2000US-0215135P.	
XX	PR	07-JUL-2000, 2000US-0216647P.	
XX	PR	07-JUL-2000, 2000US-0216880P.	
XX	PR	11-JUL-2000, 2000US-0217487P.	
XX	PR	PR 11-JUL-2000, 2000US-0217496P.	
XX	PR	14-JUL-2000, 2000US-0218290P.	
XX	PR	26-JUL-2000, 2000US-0220963P.	
XX	PR	26-JUL-2000, 2000US-0220964P.	
XX	PR	14-AUG-2000, 2000US-0224518P.	
XX	PR	14-AUG-2000, 2000US-0224519P.	
XX	PR	14-AUG-2000, 2000US-0225213P.	
XX	PR	14-AUG-2000, 2000US-0225214P.	
XX	PR	14-AUG-2000, 2000US-0225266P.	
XX	PR	14-AUG-2000, 2000US-0225267P.	
XX	PR	14-AUG-2000, 2000US-0225268P.	
XX	PR	14-AUG-2000, 2000US-0225270P.	
XX	PR	14-AUG-2000, 2000US-0225447P.	
XX	PR	14-AUG-2000, 2000US-0225757P.	
XX	PR	14-AUG-2000, 2000US-0225758P.	
XX	PR	14-AUG-2000, 2000US-0225759P.	
XX	PR	18-AUG-2000, 2000US-0226279P.	
XX	PR	22-AUG-2000, 2000US-0226568P.	
XX	PR	22-AUG-2000, 2000US-0227182P.	
XX	PR	23-AUG-2000, 2000US-0227099P.	
XX	PR	30-AUG-2000, 2000US-0228924P.	
XX	PR	01-SEP-2000, 2000US-0229287P.	
XX	PR	01-SEP-2000, 2000US-0229343P.	
XX	PR	01-SEP-2000, 2000US-0229344P.	
XX	PR	01-SEP-2000, 2000US-0229345P.	
XX	PR	05-SEP-2000, 2000US-0229509P.	
XX	PR	05-SEP-2000, 2000US-0229513P.	
XX	PR	06-SEP-2000, 2000US-0230437P.	
XX	PR	06-SEP-2000, 2000US-0230438P.	
XX	PR	08-SEP-2000, 2000US-0231242P.	
XX	PR	08-SEP-2000, 2000US-0231243P.	
XX	PR	08-SEP-2000, 2000US-0231244P.	
XX	PR	08-SEP-2000, 2000US-0231413P.	
XX	PR	08-SEP-2000, 2000US-0231414P.	
XX	PR	08-SEP-2000, 2000US-0232080P.	
XX	PR	08-SEP-2000, 2000US-0233081P.	
XX	PR	12-SEP-2000, 2000US-0233168P.	
XX	PR	14-SEP-2000, 2000US-0233397P.	
XX	PR	14-SEP-2000, 2000US-0233398P.	
XX	PR	14-SEP-2000, 2000US-0233399P.	

```

PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246529P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.

PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI, 2001-483232/52.
XX
XX Nucleic acids encoding 973 human testicular antigen polypeptides, useful
XX for preventing, diagnosing and/or treating testicular cancer.
XX
XX PS Disclosure; SEQ ID NO 2152; 766bp; English.
XX
XX CC The present invention provides the protein and coding sequences of 973
XX human testicular antigens, and fragments of their genomic sequences. The
XX sequences can be used in the treatment of cardiovascular, urinary system,
XX reproductive system, immune, respiratory, neurological and
XX gastrointestinal disorders, infections, and particularly cancer,
XX especially testicular cancers. The present sequence is a DNA encoding a
XX protein fragment of the invention
XX
XX SQ Sequence 6756 BP; 1603 A; 1601 C; 1877 G; 1675 T; 0 U; 0 Other;

Query Match 5.7%; Score 33.8; DB 4; Length 6756;
Best Local Similarity 56.9%; Pred. No. 4.3;
Matches 62; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 218 TTATATTTTCAATGATGTTTCGAGAGAGAGATGCTGTATACAGAGGCTGACTTAGA 277
DB 1737 TTATGATTTTCATTTATTTTAGAGAAAGAAATGATGAGAGCGGCTATGATTAAG 1796
QY 278 AGGTACTGACGCGGTGCGCTCTCACTAGATTCTACTAATAAATTAG 326
DB 1797 AAGAGCTCGCAGAGACATGTGACCAAAAGATATATGCTATATTAG 1845

RESULT 14
AA182806/c
ID AA182806 standard; cDNA; 360 BP.
XX
XX AC AA182806;
XX
XX DT 06-NOV-2001 (first entry)
XX
XX DE Human polynucleotide SEQ ID NO 2866.
XX
XX XX Human polynucleotide SEQ ID NO 2866.
XX
XX KM Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200164835-A2.
XX
XX PD 07-SEP-2001.
XX
XX PF 26-FEB-2001; 2001WO-US004927.
XX
XX PR 28-FEB-2000; 2000US-00515126.
XX
XX PR 18-MAY-2000; 2000US-00577409.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Tang YT, Liu C, Dymnac RT;
XX

```



DR WPI: 2001-514838/56.  
 DR P-PSDB; AAO02875.  
 XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing  
 PT and treating e.g. leukemia, inflammation and immune disorders.  
 XX  
 PS Claim 1; SEQ ID NO 2866; 1339pp + Sequence listing; English.  
 CC The invention relates to human polynucleotides (AA179941-AA193841) and  
 CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activity/inhibit activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and  
 CC inflammation. Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 360 BP; 89 A; 70 C; 106 G; 95 T; 0 U; 0 Other;  
 Query Match 5.6%; Score 33.2; DB 4; Length 360;  
 Best Local Similarity 50.6%; Pred. No. 1.6; 78; Indels 0; Gaps 0;  
 Matches 80; Conservative 0; Mismatches 78; Indels 0; Gaps 0;  
 QY 239 GAGGAGGAGATGTGTAGTACAGAGGGTACCTTAAAGTACTGACGGGTGTGCGC 298  
 Db 203 GACACGAAATATACCTGAAAGGCAAGATGCTCCCAAGTAGAGAGCCCTGGTTGGAT 144  
 QY 299 CTCTCACTAGCTTCACTAATAAATTAGAAGCTTGGCTCTACTTTCACCTGAGGCTTAA 358  
 Db 143 CCCAAGCAAGTACTTTCACCTTCCGACCATTTCCCTGTCTTTAAACGAGGCTTACA 84  
 QY 359 TTGACTTTTGTATGCGGTATATAGACAAATTACCCCA 396  
 Db 83 TTTAAATGCTATCTTATAGATCAATTACCCATTA 46  
 RESULT 15  
 AAD14902/c  
 ID AAD14902 standard; cDNA; 972 BP.  
 AC AAD14902;  
 AC 01-NOV-2001 (first entry)  
 DE Human transporter-related protein #15 cDNA.  
 DE Human transporter-related protein #15 cDNA.  
 XX Human; transporter-related protein; metabolite transporter;  
 KM organic cation transporter; multi-drug resistance; MDR; gene therapy;  
 KM cellular transporter; sodium-glucose cotransporter; diagnosis; screening;  
 KM symptomatic; phenotypic manifestation; biological condition; ss.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT 1..972  
 FT CDS /tag= a  
 FT /product= "Human transporter-related protein"  
 FT  
 XX MO200157214-A2.  
 XX  
 XX 09-AUG-2001.  
 XX  
 XX 02-FEB-2001; 2001WO-US003646.  
 XX  
 XX 03-FEB-2000; 2000US-0179973P.  
 PR 14-FEB-2000; 2000US-0182422P.  
 XX  
 XX (LEXI-) LEXICON GENETICS INC.

XX Turner CA, Mathur B, Wang X, Abuin A, Friedrich G, Zambrowicz B;  
 PI Sands AT;  
 XX  
 XX WPI: 2001-514599/56.  
 DR P-PSDB; AAE08068.  
 XX  
 PT Novel polynucleotides encoding novel human proteins with structural  
 PT similarity to cellular transporters for the diagnosis of disease and use  
 XX in gene therapy.  
 PS Disclosure; Page 53; 91pp; English.  
 CC The present sequence is a human transporter-related protein cDNA. The  
 CC human transporter-related protein share structural similarity with  
 CC mammalian metabolite or organic cation transporters, multi-drug  
 CC resistance (MDR) proteins, mammalian sodium-glucose cotransporters and  
 CC other cellular transporters. The transporter-related protein DNA may be  
 CC used for the detection of mutant sequences or inappropriately expressed  
 CC sequences for the diagnosis of disease. They may also be used to screen  
 CC for drugs effective in the treatment of the symptomatic or phenotypic  
 CC manifestations of perturbing the normal functions of the sequences of the  
 CC invention in the body. They may also be used in gene therapy for treating  
 CC biological conditions  
 XX  
 SQ Sequence 972 BP; 192 A; 257 C; 242 G; 280 T; 0 U; 1 Other;  
 Query Match 5.5%; Score 32.6; DB 4; Length 972;  
 Best Local Similarity 53.5%; Pred. No. 4.2; 59; Indels 0; Gaps 0;  
 Matches 68; Conservative 0; Mismatches 59; Indels 0; Gaps 0;  
 QY 453 TTTCTGGGTACTGCCCCGAGCTCTCTGAATTACAGCAAGTAGAAGATGTGCGAG 512  
 Db 954 TGTCTTATAGTCTCTCATTAATCTCACAAGCAAGTAAGTGAAGTGTAAACCA 895  
 QY 513 TATGACCTCTAAAAGTGAAGGTGAGTGTAAATACCACTGAGCAATCTGCCCA 572  
 Db 894 CAGCAAGCTCTCTGGGATATTAATGAGGCCACCATTTGACCAATGCGACACAGGAGCC 835  
 QY 573 GCTAGAT 579  
 Db 834 GCTGTT 828

Search completed: April 6, 2004, 23:56:11  
 Job time : 307 secs

more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0A1040CF04NP1&cluster=8230.r. Contact :  
Peng Liang Email : filiang@lifetech.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS0A1040CF04NP1.

Location/Qualifiers  
1. .1103  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1040Y107"  
/feature="CS0D1040Y107" COT 25-NORMALIZED

FEATURES  
source  
Location/Qualifiers  
1. .1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0D1072YF05"  
/feature="CS0D1072YF05" COT 25-NORMALIZED  
/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

Fri Apr 9 15:30:05 2004

us-09-613-486-14.rst

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2004, 23:23:08 ; Search time 2367 Seconds  
(without alignments)  
7531.779 Million cell updates/sec

Title: US-09-613-486-14

Perfect score: 597  
Sequence: 1 atggagctgctgcgcacag.....gtagaagggaagttatgtaa 597

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931080276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vit:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vit1:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### ALIGNMENTS

5	36.4	6.1	1201	13	EX416567	EX416567
6	35.8	6.0	507	9	AA530213	AA530213 v146b05.r
7	35.8	6.0	545	14	CB166256	CB166256 XAB60260
8	35.6	6.0	1201	13	EX395102	EX395102 BX395102
9	35.4	5.9	1201	9	AL556952	AL556952 AL556952
10	35.2	5.9	511	28	AO562700	AO562700 HS_2079.B
11	35.2	5.9	1106	29	CNS001RE	AL078654 Drosophila
12	35.2	5.9	1201	13	EX340305	EX340305 BX340305
13	35	5.9	460	9	AA712605	AA712605 31738 Lam
14	35	5.9	814	28	BH177584	BH177584 010_G_21-
15	35	5.9	814	29	CNS07K3H	AL614543 T3 end of
16	35	5.9	830	10	BP215190	BP215190 601846951
17	35	5.9	838	29	CNS0256D	AL181678 Tetradon
18	34.8	5.8	581	29	PR0014687	AL005896 F.rubripes
19	34.8	5.8	708	28	BH933072	BH933072 Ode06c10.
20	34.8	5.8	1201	13	EX394291	EX394291 BX394291
21	34.6	5.8	850	29	CNS036YG	AL230641 Tetradon
22	34.4	5.8	361	12	BI050737	BI050737 RC6-GN007
23	34.4	5.8	508	10	BE141051	BE141051 MR0-HT006
24	34.4	5.8	938	13	BO882089	BO882089 AGNCOURT
25	34.4	5.8	950	13	BP720645	BP720645 AGNCOURT
26	34.4	5.8	1201	9	AL534306	AL534306 AL534306
27	34	5.7	650	12	BG776877	BG776877 602664060
28	33.8	5.7	552	29	CB369513	CB369513 C1g7-g98-
29	33.8	5.7	657	13	BQ234878	BQ234878 hdb1b06.g
30	33.8	5.7	690	12	BG535075	BG535075 602562681
31	33.6	5.6	442	10	BF595177	BF595177 su64a08.Y
32	33.6	5.6	446	9	AI900854	AI900854 sc1f03.Y
33	33.6	5.6	506	12	BG790558	BG790558 ba674a05.
34	33.6	5.6	519	14	CA785448	CA785448 sau12b2.
35	33.6	5.6	588	13	BQ094372	BQ094372 san46c09.
36	33.6	5.6	848	29	CNS03USZ	AL261548 Tetradon
37	33.4	5.6	428	14	CB794203	CB794203 AMGNUTC:M
38	33.4	5.6	690	13	BM300017	BM300017 BM300017
39	33.4	5.6	830	28	AQ367141	AQ367141 tox00001M
40	33.4	5.6	818	29	CNS0202P	AL176254 Tetradon
41	33.4	5.6	952	12	BI079611	BI079611 602876432
42	33.4	5.6	1201	13	EX461787	EX461787 BX461787
43	33.2	5.6	503	28	AQ403529	AQ403529 HS_5055_B
44	33.2	5.6	760	28	AZ732303	AZ732303 RPCI-24-H
45	33.2	5.6	831	10	BF577284	BF577284 602135106

RESULT 1  
BE915548/c  
LOCUS  
DEFINITION BE915548 693 bp mRNA EST 29-SEP-2000  
601668490F1 NCI\_GAP\_Mam1 Mus musculus CDNA clone IMAGE:3968231 5',  
mRNA sequence.  
ACCESSION BE915548  
VERSION BE915548.1 GI:10415505  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
EST.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 693)  
REFERENCE  
NIH-MGC http://mgc.nci.nih.gov/  
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
COMMENT  
Contact: Robert Strusberg, Ph.D.

ORIGIN

Query Match 6.8%; Score 40.4; DB 10; Length 693;  
Best Local Similarity 52.4%; Pred. No. 0.27;

/clone="IMAGE:3968231"  
/tissue\_type="tumor, biopsy sample"  
/dev\_stage="10 months, virgin"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Mam1"  
/note="Organ: mammary; Vector: PCMV-SPORT6; Site: 1; Salt; Site 2: Not; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

Fri Apr 9 15:30:05 2004

us-09-613-486-14.rst

Page 4

SOURCE ORGANISM

Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1201)  
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifestech.com URL : http://fulllength.invitrogen.com/InvitrogenCorporation 1600 Faraday Avenue Genoscope sequence ID : CSODA008D100P1.

FEATURES

source

Location/Qualifiers  
1..1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODA008YH20"  
/tissue\_type="NEUROBLASTOMA"  
/clone\_lib="Homo sapiens NEUROBLASTOMA"  
/note="Vector: PCMVSPORT 6; 1st strand cDNA was primed with a NotI-Oligo (dt) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the PCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 6.1%; Score 36.4; DB 13; Length 1201;  
Best Local Similarity 15.3%; Pred. No. 7;  
Matches 41; Conservative 88; Mismatches 138; Indels 1; Gaps 1;

98 AAAAAAGTTGGTCAGAAAGGGGCTCTAAGAGGATATAGAGTGGTTCAGTAC 157  
Db ANNNNAANNKGGGGTGCAGAAAGNNNNNNNNAANGGGGNNNGGNNNGNNNTNN 534  
QY TCCTTTAGCACTCGCGGCAAGAACACAGCTCTCTAAGATTCAAGCGGCAAGTTCAGACG 217  
Db NNN 594  
QY TTAATATTTTCAATAGTTTGGAGAGAGAAATGTGTAGTACAGAGGTGACCTTATAGA 277  
Db CTAKKTKKNNCMMATKTKKKKTTKTKKKKTAMKMMCMAMMSMGYKGGSVSCAMVM 654  
QY AGGTAAGTCAAGCGGTGTGGCGCTCTCACTAGCTTCACTAATAAATTAGACGTTGCGTC 337  
Db CGGWSMDMKCMKTKGKNNMMCKKCMKCKKGGHGTGM-NNMCTGHHMMYKKTACKB 713  
QY GTACTTCACTGAGGCTTACGTTGACTT 365  
Db WTKYKCTMCTNKTKTCTMCKKVKKK 741

RESULT 6  
AA530213 507 bp mRNA linear EST 22-JUL-1997  
LOCUS v466h05.r1 StrataGene mouse skin (#937313) Mus musculus cDNA clone  
DEFINITION IMAGE:932121 5', mRNA sequence.  
ACCESSION AA530213  
VERSION AA530213.1 GI:2272919

Email: cgsaps-remail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LUNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
Contact: amgdc@bcm.tmc.edu  
Gunnar, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Louisged, H., Kowals, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Navevati, A.N., Gibbs, R.A.

TITLE JOURNAL COMMENT

Waterston, R.  
The WashU-HMI Mouse EST Project  
Unpublished (1996)  
Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@waterston.wustl.edu  
This clone is available royalty-free through LUNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -28m3 rev1 ET from Amersham  
High quality sequence stop: 162.

FEATURES

source

Location/Qualifiers  
1..507  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE:932121"  
/sex="females"  
/tissue\_type="whole skin"  
/dev\_stage="11 weeks old"  
/lab\_host="SOLR (kanamycin resistant)"  
/clone\_lib="Stratagene mouse skin (#937313)"  
/note="Organ: skin; Vector: pBluescript SK-; Site 1: Scori; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Whole skin from 11 week old C57BL/6 female mice. Average insert size: 1.0 kb. Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGG 3' -3' adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3' "

ORIGIN

Query Match 6.0%; Score 35.8; DB 9; Length 507;  
Best Local Similarity 51.6%; Pred. No. 6.7;  
Matches 82; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 30 CAACCTGGTATACCGACGCTCTAGTCTAATAGTGTCAACAAGAGCTTTATCTGC 89  
Db CATCTCGTGGGTGTCTCTAGCTTCAAGTCAAGTGTGCTTCAAGATGA 359  
QY TGAAGTGAATAAAGTTGGTGCAGAAAGGGGCTCTAAGAGGATATAGAGTGTGTT 149  
Db TCAAGCGAAGAGCTGTTCCCGAGTAAAGGCACTCGATGTAGAGAAAGAGTGT 419  
QY 150 CGGTCTACTCTTTAAGCACTGCGCGGCAAGAACACGTC 188  
Db CAGCGGTGACCATCATGACCATCATATGACGATCAGACGTC 458

RESULT 7  
CB166256 545 bp mRNA linear EST 30-JAN-2003  
LOCUS CB166256  
DEFINITION XAB602600277.R1 CSEBFX03 kidney Bos taurus cDNA, mRNA sequence.  
ACCESSION CB166256  
VERSION CB166256.1 GI:28152381  
KEYWORDS EST.  
SOURCE Bos taurus (cow)  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Tel: 9798452616  
Fax: 9798456970  
Email: david.adelson@amu.edu.  
Location/Qualifiers

## FEATURES

## source

1..545  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"  
/tissue\_type="kidney"  
/clone\_lib="CEGXP.L03 kidney"  
/note="Organ: kidney; Vector: pBluescript SK+; Site\_1: NotI; Site\_2: EcoRI; sequence 5' of the insert (5'-NNN...NNNinsert)  
GCGAATGAGCTCCACCGCGGTGGCGCGCGCGCTCGAG. Sequence 3' of the inserts (AAGATTGATATCAAGCTTATGACCTGACCTCGAG. non-normalized library, sequenced 3' with M13R primer."

## ORIGIN

Query Match 6.0%; Score 35.8; DB 14; Length 545;  
Best Local Similarity 52.3%; Pred. No. 7;  
Matches 79; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 430 GAAGATTGCTACTAGCTGCAGATTTCTGGTACTTCCCGAAGCTCTCTGAATTACAG 489  
DB 94 GATGCTGATAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 153  
QY 490 CAAGTGAAGAGATGTTCCGAGATGATGATGATGATGATGATGATGATGATGATGAT 549  
DB 154 CAAGAAGAGATCATCTGCTGTTGTTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGTCA 213  
QY 550 ACACGAGGAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 580  
DB 214 ACAGATTGAGGATCAAGCTGACCTAGCAA 244

RESULT 8  
BX395102 1201 bp mRNA linear EST 13-MAY-2003  
LOCUS BX395102 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
DEFINITION cDNA clone CS0DC029YK20 5-PRIME, mRNA sequence.

ACCESSION BX395102.1 GI:30616562  
VERSION BX395102.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1201)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
BP 191 91006 EVRY cedex - France  
Email: segr@genoscope.cns.fr, Web: www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 2431.f. For more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DC029BP100P1&cluster=2431.f. Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID: CS0DC029BP100P1.

## FEATURES

## source

1..1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DC029YK20"  
/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 6.0%; Score 35.6; DB 13; Length 1201;  
Best Local Similarity 45.2%; Pred. No. 13;  
Matches 57; Conservative 17; Mismatches 52; Indels 0; Gaps 0;

QY 211 TCAGACGTTATATTTCAATAGTTTCGAGAGAGAGATGTGTAGTAACAGAGGTGAC 270  
DB 1071 TCARAGTTATATTTGTAATATAATATAATATGAKATGCGYKTTATAGGKTGTT 1110  
QY 271 CTTAAGAAGTACTGAGCGGTGTCGCTCTCACTAGTTCCTCACTAATTAAGTAAAGC 330  
DB 1131 CTTGAAACCHYTWCTTGTGTCWGGGCTTAAVCTCCCAATTAAGCATTCYATGG 1150  
QY 331 TTGCGT 336  
DB 1191 TTGKMT 1196

RESULT 9  
AL556952 1201 bp mRNA linear EST 31-MAY-2003  
LOCUS AL556952 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA  
DEFINITION clone CS0DH005YA07 3-PRIME, mRNA sequence.

ACCESSION AL556952.2 GI:31278752  
VERSION AL556952.2  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1201)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On Feb 15, 2001 this sequence version replaced gi:12300090.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segr@genoscope.cns.fr, Web: www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7763.r. For more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DH005AA04NP1&cluster=7763.r. Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID: CS0DH005AA04NP1.

## FEATURES

## source

1..1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DH005YA07"  
/tissue\_type="T CELLS (JURKAT CELL LINE)"  
/cell\_line="JURKAT CELL LINE"  
/clone\_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

## ORIGIN

Query Match 5.9%; Score 35.4; DB 9; Length 1201;  
Best Local Similarity 42.3%; Pred. No. 15;  
Matches 85; Conservative 18; Mismatches 98; Indels 0; Gaps 0;

QY 70 GACAAGACCTTTATCTGCTGAAGTGAATAAATGTTGGCGAAGAGGGCTCTCTAC 129  
DB 1 KATTAATVACTTATGTTGTAACANAGCTGAGCATGATANGCWCACGAGACTCTGT 60

QY 130 GAGGATATAGAAAGTGTTCCTTACCTTACGACCTCGCGGAGAACCAACGCTC 189  
 Db 61 GTGGGTTTATGAGGAGTGTGATTAACCTCGTGNHGBGGGCTGTGTGATVATACAGCTTC 120  
 QY 190 CTTAAGCTTACGCGGAGATTCAGACCTTATATTTTCAATAGTTTCGAGAGAGAAAT 249  
 Db 121 CTTCCWDBBGTGGGGGTGAGGTAGCTGTGTGTTAKDTATKGBATCAKKGTTGGCTT 180  
 QY 250 GTGTACTATACAGAGGTGAC 270  
 Db 181 GGBAAGKGGCCGAGGGTGC 201

RESULT 10  
 A0562700/c  
 LOCUS 511 bp DNA linear GSS 29-MAY-1999  
 DEFINITION HS\_2079\_B1\_B05\_T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2079 Col=9 Row=D, genomic survey  
 sequence.  
 ACCESSION A0562700  
 VERSION A0562700.1 GI:4922171  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 511)  
 AUTHORS Mahairas G.G., Wallace J.C., Smith K., Swartzell S., Holzman T., Keller A., Shaker R., Furlong J., Young J., Zhao S., Adams M.D. and Hood L.  
 COMMENT Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
 JOURNAL 99380589  
 MEDLINE 1049764  
 PUBMED  
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887  
 Email: jwallace@u.washington.edu  
 Clones may be purchased from Research Genetics (info@resgen.com).  
 BAC end Web Server: http://www.htsc.washington.edu  
 Plate: 2079 row: D column: 9  
 Seq primer: T7  
 Class: BAC ends  
 High quality sequence stop: 511.  
 Location/Qualifiers  
 1..511  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="Plate=2079 Col=9 Row=D"  
 /sex="male"  
 /clone\_lib="CIT Approved Human Genomic Sperm Library D"  
 /note="Organ: sperm; Vector: pBelbac11; BAC clones in E-Coli DH10B"

ORIGIN  
 Query Match 5.9%; Score 35.2; DB 28; Length 511;  
 Best Local Similarity 48.3%; Pred. No. 10;  
 Matches 97; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

FEATURES  
 source  
 1..511  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="Plate=2079 Col=9 Row=D"  
 /sex="male"  
 /clone\_lib="CIT Approved Human Genomic Sperm Library D"  
 /note="Organ: sperm; Vector: pBelbac11; BAC clones in E-Coli DH10B"

QY 211 TCAGACGTATATTTTCAATAGTTTCGAGAGAGAAATGTGTAGTACAGAGGTGAC 270  
 Db 101 GCAACATAGAAATGCAAGAGAGCTGCTGCTCAGACCTGATCGCGGAGGTGCTA 42  
 QY 271 CTTAAGAGGTACTGACGG 291  
 Db 41 AGTATCAGGAGTGTGACGG 21

RESULT 11  
 CDS00LRE 1106 bp DNA linear GSS 14-JUN-1999  
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC: BACR48P05 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
 AL078654  
 AL078654.1 GI:5101944  
 GSS.  
 ACCESSION Drosophila melanogaster (fruit fly)  
 Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 REFERENCE 1 (bases 1 to 1106)  
 Genoscope.  
 DIRECT SUBMISSION  
 Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
 Web : www.genoscope.cns.fr  
 The BACP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mammaster in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the library P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).  
 Location/Qualifiers  
 1..1106  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /clone="BACR48P05"  
 /clone\_lib="RPCI-98"  
 /note="end : T7"

ORIGIN  
 Query Match 5.9%; Score 35.2; DB 29; Length 1106;  
 Best Local Similarity 15.1%; Pred. No. 16;  
 Matches 47; Conservative 132; Mismatches 133; Indels 0; Gaps 0;

FEATURES  
 source  
 1..1106  
 /organism="Drosophila melanogaster"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7227"  
 /clone="BACR48P05"  
 /clone\_lib="RPCI-98"  
 /note="end : T7"

QY 371 TCGCGTATAGCACAATTACCCCAAGCCGCGGGAATTGGGATTCACAGTG 430  
 Db 954 WTTTIDWVWVRGVBKDVAVVMMWBCWMAWNGVGVHAWTDEKSSHYBNBDIY 1013  
 QY 431 AAGATTGCTACT 442  
 Db 1014 SWVANHTRKANB 1025

RESULT 12  
 EX340305 1201 bp mRNA linear EST 02-MAY-2003  
 LOCUS EX340305 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CSOD1085Y110 5-PRIME, mRNA sequence.  
 ACCESSION EX340305  
 VERSION EX340305.1 GI:30333886  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 COMMENT Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1201)  
 Li, W.B., Gruber, C., Jesse, J. and Polyes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr Web: www.genoscope.cns.fr  
 Library was constructed by life technologies, a division of Invitrogen. This sequence belongs to sequence cluster 10798.f For more information about this cluster, see  
 http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSOD1085BEO5QPI&cluster=10798.f. Contact: Feng liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Reddick Avenue genoscope sequence ID: CSOD1085BEO5QPI.  
 Location/Qualifiers  
 1. 1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CSOD1085Y110"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_id="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
 Query Match 5.9%; Score 35.2; DB 13; Length 1201;  
 Best Local Similarity 46.4%; Pred. No. 17;  
 Matches 64; Conservative 13; Mismatches 61; Indels 0; Gaps 0;

QY 449 CAGATTTCGGAGTCTGCCGAGCTCTGTGATTAACAGCAAGTAGAGATGTTG 508  
 Db 1001 CTGAGCTTTCGCCACCCCTCAGGAGAGTCMTTGGRAACAAMATTTGGCAGRGCCCT 1060  
 QY 509 CGAGTATGTCGCTCTAAAACCTGAGAGTGAGTGTAAATACCAAGTAGCAATCTGC 568  
 Db 1061 CAGAGTTCCACACTCTACTCTCCAGTGCTTTTTCMTTCAACAATCACAAGAGT 1120  
 QY 569 GTGAGTACGAGTAGAGG 586  
 Db 1121 TWWGATTTTCTCATSGS 1138

RESULT 13  
 AA712605 460 bp mRNA linear EST 24-DEC-1997  
 LOCUS AA712605  
 DEFINITION 31738 Lambda-PRU2 Arabidopsis thaliana cDNA clone 176N177, mRNA

sequence.  
 ACCESSION AA712605  
 VERSION AA712605.1 GI:2722522  
 KEYWORDS EST.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 COMMENT Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 460)  
 Newman, T., deBunijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomasow, M., Retzel, E. and Somerville, C.  
 Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones  
 Plant Physiol. 106, 1241-1255 (1994)  
 95148729  
 MEDLINE  
 JOURNAL  
 PubMed  
 COMMENT Contact: Thomas Newman  
 MSU-DOE Plant Research Laboratory  
 Michigan State University  
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI  
 Tel: 517-353-0854  
 Fax: 517-353-9168  
 Email: 2213@cmeibm.c1.msu.edu  
 Seq primer: T7 dye primer.  
 Location/Qualifiers  
 1. 460  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /strain="var Columbia"  
 /db\_xref="taxon:3702"  
 /clone="176N177"  
 /clone\_id="Lambda-PRU2"  
 /note="Vector: lambda zip-lox; Site 1: Sal; Site 2: Not; lambda PRU2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRU's lambda zip-lox. The cDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."

ORIGIN  
 Query Match 5.9%; Score 35; DB 9; Length 460;  
 Best Local Similarity 50.7%; Pred. No. 11;  
 Matches 71; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 445 GGTGAGATTTCTGGAGTCTGCCGAGCTCTGTGATTAACAGCAAGTAGAGATG 504  
 Db 222 GATTAGCGGCTCTTTCTTTCATCAGAGTTTATGCTTCAAGTAATTAATATC 351  
 QY 505 TTGGAGATATGATGCTCTAAAACCTGAGAGTGAGTGTAAATACCAAGTAGCAAT 564  
 Db 352 ATCACTGTTGTGTTGCTATGAAAANAANGTGNGTTGAAATAAAGAGGAGCGCT 411  
 QY 565 CTGCGTACGAGTAGAGAG 584  
 Db 412 NGGTTTAACTTTGGGAGAG 431

RESULT 14  
 BH177584 814 bp DNA linear GSS 19-OCT-2001  
 LOCUS BH177584/c  
 DEFINITION 010.G.21-rev SmBAC1 Schistosoma mansoni genomic survey sequence.  
 ACCESSION BH177584  
 VERSION BH177584.1 GI:16276482  
 KEYWORDS GSS.  
 SOURCE Schistosoma mansoni

REFERENCE	TITLE	JOURNAL	MEDLINE	COMMENT
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
Other_GSSs: 010_G_21-21				
Contact: Pierce RJ				
INSERM U 167				
Institut Pasteur de Lille				
1 rue du Professeur A. Calmette, 59019-Lille, France				
Tel: (33) (0)3 20877783				
Fax: (33) (0)3 20877888				
Email: Raymond.Pierce@pasteur-lille.fr				
CNS sequencing ID=DC00A010AD11BP1				
Plate: 010 row: G column: 21				
Seq primer: M13 reverse primer				
Class: BAC ends				
High quality sequence stop: 814.				
Location/Qualifiers				
1..814				
/organism="Schistosoma mansoni"				
/mol_type="genomic DNA"				
/strain="Puerto-Rican"				
/db_xref="taxon:6183"				
/clone="010G21"				
/sex="mixed"				
/dev_stage=" cercariae"				
/lab_host="Biomphalaria glabrata"				
/clone_1ib="SmbAC1"				
/note="Vector: pBel0BAC 11; Site 1: Hind III. Partially Hind III digested and size-selected S. mansoni cercarial DNA was ligated into Hind III digested pBel0BAC 11 vector and used to transform E. coli DH10B. The complete library contains 23808 clones from 4 independent sizing-ligation-transformations. Average insert size ranges from 70-127 kb and genome coverage is 7.9-fold."				
ORIGIN				
Query Match	5.9%;	Score 35;	DB 28;	Length 814;
Best Local Similarity	63.9%;	Pred. No. 16;		
Matches	53;	Conservative	0;	Mismatches 30; Indels 0; Gaps 0
CY	193	AAGTTTCACGCCGCGCATTCAGACGCTTATTTCAATAGTTTCGAGAGGAGTGTG	252	
DB	388	AAATCTCACTTGTCTCAATCAGTAGTTATAGTCTATCGAATTCGGCAGAAATCTGTA	329	
CY	253	GTATTAACAGAGGTGACCTTAA	275	
DB	328	GTATTAACAACGGTGGCAATGA	306	
RESULT 15				
CNS07K3H/c	814 bp	DNA	linear	GSS 30-NOV-2001
LOCUS	T3 end of clone 010AD11	of library SmbAC1	from strain	Puerto-Rican
DEFINITION	of Schistosoma mansoni,	genomic survey	sequence.	
ACCESSION	AF614543			
VERSION	AL614543.1	GI:16027767		
KEYWORDS	GSS.			
SOURCE	Schistosoma mansoni			
ORGANISM	Schistosoma mansoni			
REFERENCE	Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;			
AUTHORS	Williams,D.L., Johnston,D., Loverde,P.T. and Le Paslier,D.			
TITLE	Construction and characterization of a Schistosoma mansoni			
JOURNAL	bacterial artificial chromosome library			
MEDLINE	Genomics 65 (2), 87-94 (2000)			
PUBMED	20247247			
Other_GSSs: 010_G_21-21				
Contact: Pierce RJ				
INSERM U 167				
Institut Pasteur de Lille				
1 rue du Professeur A. Calmette, 59019-Lille, France				
Tel: (33) (0)3 20877783				
Fax: (33) (0)3 20877888				
Email: Raymond.Pierce@pasteur-lille.fr				
CNS sequencing ID=DC00A010AD11BP1				
Plate: 010 row: G column: 21				
Seq primer: M13 reverse primer				
Class: BAC ends				
High quality sequence stop: 814.				
Location/Qualifiers				
1..814				
/organism="Schistosoma mansoni"				
/mol_type="genomic DNA"				
/strain="Puerto-Rican"				
/db_xref="taxon:6183"				
/clone="010G21"				
/sex="mixed"				
/dev_stage=" cercariae"				
/lab_host="Biomphalaria glabrata"				
/clone_1ib="SmbAC1"				
/note="Vector: pBel0BAC 11; Site 1: Hind III. Partially Hind III digested and size-selected S. mansoni cercarial DNA was ligated into Hind III digested pBel0BAC 11 vector and used to transform E. coli DH10B. The complete library contains 23808 clones from 4 independent sizing-ligation-transformations. Average insert size ranges from 70-127 kb and genome coverage is 7.9-fold."				

```

JOURNAL      Bacterial artificial chromosome library
Genomics    65 (2), 87-94 (2000)
MEDLINE      20247247
PubMed       10783255
REFERENCE    2 (bases 1 to 814)
AUTHORS      Genoscope.
TITLE        Direct SubMISSION
JOURNAL      Submitted (05-OCT-2001) Genoscope - Centre National de Sequencage
BP 191, 91006 EVRY cedex - FRANCE (E-mail : segret@genoscope.cns.fr)

COMMENT
  - Web : www.genoscope.cns.fr)
  - Partially Hind III digested and size-selected S. mansoni cercarial
    DNA was ligated into Hind III digested pBelosAC 11 vector and used
    to transform E. coli DH103. The complete library contains 23808
    clones from 4 independent sizing-ligation-transformations. Average
    insert size ranges from 70-127 kb and genome coverage is 7.9-fold.

FEATURES
  source
    Location/Qualifiers
      1..814
        /organism="Schistosoma mansoni"
        /mol_type="genomic DNA"
        /strain="Puerto-Rican"
        /db_xref="taxon:6183"
        /clone="010MD11"
        /clone_11b="SmbAC1"
        /note="end : T3"

ORIGIN
Query Match      5.9%; Score 35; DB 29; Length 814;
Best Local Similarity 63.9%; Pred. No. 16;
Matches 53; Conservative 0; Mismatches 30; Indels 0; Gaps 0

D1 193 AAGGTTACGCGCGCAAGATTACAGACGTTATTTTCAATACGTTTCGAGAGAGAAATGTG 252
D2 388 AATTCCTCACTTGCTCAATCGACGATGATTAATAGTCATCGAATTCGCGCGCAAGATCTGTA 329
D3 253 GTAGTAAACAGAGGGGTACCTTAA 275
D4 328 GTATTACAAACGGTGGCAATTA 306

Search completed: April 7, 2004, 01:14:08
Job time : 2372 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 5, 2004, 17:46:17 ; Search time 60 Seconds  
(without alignments)  
932.407 Million cell updates/sec

Title: US-09-613-486-15

Perfect score: 991  
Sequence: 1 MELMSDNLNLTVDASSL.....GGVNTPVSNLRQLGRREVM 198

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991	100.0	199	2	AAW73482 Grapevine
2	286	28.9	204	2	AAW72682 Sugar bee
3	89.5	9.0	499	5	ABG3272 C. albica
4	86.5	8.7	328	5	ABBS1963 Herbicida
5	86	8.7	295	3	AAAG3167 Arabidops
6	86	8.7	295	3	AAAG20735 Arabidops
7	86	8.7	360	3	AAAG20734 Arabidops
8	86	8.7	360	3	AAAG3166 Arabidops
9	86	8.7	360	5	ABBS1962 Herbicida
10	86	8.7	366	3	AAAG3165 Arabidops
11	86	8.7	366	3	AAAG20733 Arabidops
12	86	8.7	447	4	AAAB6072 Putative
13	85	8.6	599	2	AAW17788 Phage abo
14	84.5	8.5	269	4	AAV97690 GRNAV-5 C
15	83.5	8.4	422	6	ABU26015 Protein e
16	82.5	8.3	591	6	ABU43731 Protein e
17	81.5	8.2	402	3	AAAB4800 Erysipelo
18	81.5	8.2	402	5	AAAB0869 Erysipelo
19	81.5	8.2	402	5	ABBO7787 E. thusiop
20	81.5	8.2	626	3	AAV95782 Erysipelo
21	81.5	8.2	813	6	ABU19284 Protein e
22	80	8.1	555	5	ABBS4251 Lactococc
23	80	8.1	711	5	ABP29955 Streptoco
24	80	8.1	711	6	ABU46760 Protein e
25	80	8.1	729	5	ABP25672 Streptoco

26	79.5	8.0	770	2	AAW26584 Rat hemid
27	79	8.0	263	4	AAU49697 Propionib
28	79	8.0	263	6	ABM46216 Propionib
29	79	8.0	451	4	AAU35128 Enterococ
30	79	8.0	451	6	ABU29218 Protein e
31	79	8.0	477	2	AAW21643 Grapevine
32	79	8.0	478	6	ABU35792 Protein e
33	79	8.0	782	4	ABBS57835 Drosophil
34	78.5	7.9	208	5	ABBS54384 Lactococc
35	77	7.8	696	5	ABP73971 Candida a
36	77	7.8	865	6	ABU22392 Protein e
37	76.5	7.7	200	4	AAU30591 Novel hum
38	76	7.7	223	2	AAW73481 Grapevine
39	75.5	7.6	509	2	AAW47583 NAHD oxid
40	75.5	7.6	606	2	AAV43219 E. thusiop
41	75	7.6	358	5	ABBS1943 Herbicida
42	75	7.6	404	5	ABP25433 Streptoco
43	75	7.6	883	4	AAW51651 Synchoco
44	75	7.6	26926	4	AAU05396 Human tit
45	75	7.6	31267	6	ABG74786 Human RGS

#### ALIGNMENTS

RESULT 1  
AAW73482  
ID AAW73482 standard; protein, 198 AA.  
AC AAW73482;  
XX  
XX  
DT 27-AUG-2003 (revised)  
DT 29-MAR-1998 (first entry)  
XX  
XX  
DE Grapevine leafroll virus type 2 coat protein.  
XX  
XX  
KM GRLAV-2; closterovirus; grape; tobacco; transgenic plant;  
KM disease resistance; virus resistance; beet yellows virus; cristeza virus;  
KM coat protein.  
XX  
XX  
OS Grapevine leafroll virus.  
OS  
PN WO9853055-A1  
XX  
XX  
PD 26-NOV-1998.  
XX  
XX  
PF 20-MAY-1998; 98MO-US010313.  
XX  
XX  
PR 20-MAY-1997; 97US-0047194P.  
XX  
XX  
PA (CORR ) CORNELL RES FOUND INC.  
XX  
XX  
PI Zhu H, Ling K, Gonsalves D;  
XX  
XX  
DR WPI. 1999-045307/04.  
DR N-PSDB; AAV08670.  
XX  
XX  
PT Grapevine leafroll virus (type 2) proteins and polypeptides - and  
PT encoding DNA, useful e.g. to impart grapevine leafroll resistance to  
PT grape and tobacco plants and detect grapevine leafroll virus.  
XX  
XX  
PS Claim 12; Page 44-45; 151pp; English.  
XX  
XX  
CC This is the amino acid sequence of a 22 kDa coat protein that is encoded  
CC by open reading frame ORF6 (see AAV08870) of grapevine leafroll virus  
CC type 2 (GRLAV-2) RNA (see AAV08874). The GRLAV-2 genome includes 9 open  
CC reading frames (see AAV08864-72) for a polyprotein, an RNA-dependent RNA  
CC polymerase, heat shock proteins, coat proteins and proteins of unknown  
CC function (see AAW73476-84). These can be used to produce antibodies  
CC useful for detecting GRLAV in samples e.g. by ELISA (claimed). The  
CC nucleic acid molecules can be used to produce probes and primers for such  
CC detection, and to transform host cells (especially Agrobacterium vitis,  
CC Agrobacterium tumefaciens, grape, citrus, beet or tobacco cells) and



CC produce transgenic plants (claimed). They can be used to impart GLRAV-2  
 CC resistance to Vitis scion or rootstock cultivars or Nicotiana cultivars  
 CC (claimed). Because extensive similarity exists between hsp70-related  
 CC sequence regions of GLRAV-2 and other closteroviruses, the DNA may also  
 CC be used to impart beet yellows virus resistance to beet cultivars or  
 CC tristea virus resistance to citrus scion cultivar/rootstock cultivars  
 CC (claimed). (Updated on 27-Aug-2003 to correct OS field.)  
 CC

XX Sequence 198 AA;

Query Match 100.0%; Score 991; DB 2; Length 198;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-100;  
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELMSDSNLSNLTVDASSLNGVDKLLSAEVEKMLVQKAPNEGIEVFGLLYALAAAR 60  
 DB 1 MELMSDSNLSNLTVDASSLNGVDKLLSAEVEKMLVQKAPNEGIEVFGLLYALAAAR 60  
 QY 61 TTSPKVOQADSDVIFSNFSGERNVVTBGLDKVLDGCAPLTFRTNKLRTFGSTFEAYV 120  
 DB 61 TTSPKVOQADSDVIFSNFSGERNVVTBGLDKVLDGCAPLTFRTNKLRTFGSTFEAYV 120  
 QY 121 DFCIAYKHKLPOLNAAAEELGIPADSYLAADFLGTCPKSELQOSRMPASMYALKTEGG 180  
 DB 121 DFCIAYKHKLPOLNAAAEELGIPADSYLAADFLGTCPKSELQOSRMPASMYALKTEGG 180  
 QY 181 VVNTPVSNLRQLGRREV 198  
 DB 181 VVNTPVSNLRQLGRREV 198

RESULT 2

AA72682 standard; protein; 204 AA.

AC AAR72682;  
 DT 16-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 02-NOV-1995 (first entry)

XX Sugar beet yellows virus capsid protein.  
 XX Sugar beet yellows virus capsid protein; transgenic plant.  
 XX Beet yellows virus.  
 XX RU2017820-C1.  
 XX 15-AUG-1994.  
 XX 27-JUN-1991; 91SU-04950054.  
 XX 27-JUN-1991; 91SU-04950054.  
 XX (IMMUNO) IMMUNOBIOTECH INST.  
 XX (BIOT) BIOYRCHN INST CO LTD.  
 XX Arganovskii AA, Bolko VP, Karasev AV;  
 XX MPI; 1995-113715/15.  
 XX N-PSDB; AAQ87653.  
 PT Sugar beet yellows virus cDNA fragment encoding capsid protein - useful  
 PT for production of virus-resistant transgenic plants.  
 XX Claim 1; Col 7-10; 5pp; Russian.

CC The amino acid sequence of the novel sugar beet yellows virus (SBYV)  
 CC capsid protein. The protein has mol. wt 22.2 kD. The corresponding gene  
 CC was obtained from reverse transcribed RNA isolated from purified SBYV.  
 CC The fragment is useful for the production of virus resistant transgenic  
 CC plants by genetic engineering methods. (Updated on 25-MAR-2003 to correct

CC PN field.) (Updated on 25-MAR-2003 to correct PF field.) (Updated on 16-  
 CC OCT-2003 to standardise OS field)

XX Sequence 204 AA;

Query Match 28.9%; Score 286; DB 2; Length 204;  
 Best Local Similarity 37.6%; Pred. No. 8.5e-23;  
 Matches 71; Conservative 29; Mismatches 87; Indels 2; Gaps 2;

QY 7 SNLSNLTVDASSLNGVDKLLSAEVEKMLVQKAPNEGIEVFGLLYALAAFTSPKV 66  
 DB 11 AFFENVSLADQCTLHGDCDCKLRKNFEBCCLKGVPEIDNLGIALGLCYSCATIGTSNKV 70  
 QY 67 QPADSDVIFSNF-GERNVVTBGLDKVLDGCAPLTFRTNKLRTFGSTFEAYVDFCIA 125  
 DB 71 NVQPTSTTKASFQGGKELVTHGELNSFLSQKLSKPKKLCFCRTFQKVIISLKE 130  
 QY 126 YHKLPOLNAAAEELGIPADSYLAADFLGTCPKSELQOSRMPASMYALKTEGGVNT 185  
 DB 131 YRGKLPPIAANRHGLPADHYLAADFISTSTELDLQOSRLLAENATHTERS-SESP 189  
 QY 186 VSNLRQLGR 194  
 DB 190 VSNLRQLGR 198

RESULT 3

ABG93273 standard; protein; 499 AA.

AC ABG93273;  
 DT 21-NOV-2002 (first entry)

XX C. albicans Bax-associated protein fragment SEQ ID 504.

XX Bax; Bax-resistance; cytosolic; fungicide; immunosuppressive; virucide;  
 XX vasotropic; vaccine; gene therapy; proliferative disorder; cancer;  
 XX apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;  
 XX neurodegeneration; cell death.

XX Candida albicans.

XX WO200264766-A2.

XX 22-AUG-2002.

XX 21-DEC-2001; 2001WO-EP015398.

XX 22-DEC-2000; 2000EP-00870318.

XX 04-JAN-2001; 2001EP-00870002.

XX 09-JAN-2001; 2001EP-00870003.

XX (JANC) JANSSEN PHARM NV.

XX Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;  
 XX MPI; 2002-667002/71.  
 XX N-PSDB; ABQ6539.

XX New isolated nucleic acid representing a synthetic BAX-gene, useful as  
 XX medicament for treating, preventing and/or alleviating yeast or fungal  
 XX infections or proliferative disorders, or for preventing apoptosis in  
 XX certain diseases.

XX Claim 36; Fig 2; 344pp; English.

CC This invention describes a novel nucleic acid representing a synthetic  
 CC Bax gene. The Bax gene of the invention is useful for identifying Bax-  
 CC resistant yeast or fungi, identifying, or obtaining and identifying  
 CC Candida spp. sequences that are differentially expressed in a pathway  
 CC eventually leading to programmed cell death or identifying inhibitors or  
 CC inhibitor sequences of Bax-induced cell death. The products of the

CC invention have cytostatic, fungicide, immunosuppressive, virucide and  
 CC vasotropic activity and can be used in vaccines or for gene therapy. The  
 CC isolated nucleic acids, polypeptides, pharmaceutical compositions,  
 CC antitense molecules and antibodies are useful as medicaments or in  
 CC preparing a medicament for treating, preventing and/or alleviating  
 CC diseases associated with yeast or fungi or proliferative disorders, such  
 CC as cancer, or for preventing apoptosis in certain diseases. The compounds  
 CC or polypeptides, or the genetically modified organism are useful for  
 CC preparing a medicament for modifying the endogenous flora of humans and  
 CC other mammals. The vaccine is useful for immunising against yeast or  
 CC fungal infections. Apoptosis-related diseases include autoimmune diseases,  
 CC ischaemia, diseases related with viral infections or neurodegenerations.  
 CC This sequence represents a polypeptide associated with the Bax gene  
 CC described in the disclosure of the invention

XX Sequence 499 AA;

Query Match 9.0%; Score 89.5; DB 5; Length 499;  
 Best Local Similarity 25.1%; Pred. No. 1.2;  
 Matches 53; Conservative 39; Mismatches 76; Indels 43; Gaps 12;

QY 3 LMSDSNLSNLTVPASS-LNGVDKLLSA-EVEKIVQKAPNEGIEVFGLLVLAAR 60  
 DB 173 LMSWKGPLVATGTTVLTAESTPLSALYLSQLVEAGMPGVNIVSGFATKGAAL 232  
 QY 61 TTSPKVRADSDVIFSNFGERNVV---TEGDLKV---LDGCAPIRTFNKLTFTGRT 114  
 DB 233 AKHKIKR---VAFTGSTATGKIIMLAESNLKVTLELGGSPNIVFND--ADLDKT 286  
 QY 115 FTEAYV-----DFCTA-----YKHKLPLNAAAE---LGTF-AEDSYLAADFL 153  
 DB 287 IONLIVIFYNSSGVCCAGSRLLIQSGVYDQVVEKFEAAESYKGVNPFDEDTFMA--- 343  
 QY 154 GTCPELSELOOSRKMFAFWALKTGEGVNT 184  
 DB 344 ---QVSDVQLS-KILKVESGKSGGATVTT 369

RESULT 4  
 ABB91963  
 ID ABB91963 standard; protein, 328 AA.

XX ABB91963;

XX 31-MAY-2002 (first entry)

XX Herbicidally active polypeptide SEQ ID NO 1174.

XX Herbicidally plant; agriculture; herbicide.

XX Arabidopsis thaliana.

XX WO200210210-A2.

XX 07-FEB-2002.

XX 28-AUG-2001; 2001WO-EP009892.

XX 28-AUG-2001; 2001WO-EP009892.

XX (FARB ) BAYER AG.

XX Tietjen K, Weidner M;

XX WPT 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,  
 PT comprising aligning and comparing nucleic acid or amino acid sequences  
 PT from plant with nucleic acid or amino acid sequences from non-plant  
 PT organisms.

PS Claim 5; SEQ ID NO 1174; 261pp + Sequence Listing; English.

CC The invention relates to identifying target proteins (ABB90790-ABB94016)  
 CC for herbicidally active compounds, comprising aligning and comparing  
 CC nucleic acid or amino acid sequences from plant with nucleic acid or  
 CC amino acid sequences from non-plant organisms using suitable search  
 CC parameters, where plant sequences having an E-value greater by a factor  
 CC of 3 than the E-value of most similar non-plant sequences are selected.  
 CC The polypeptides or nucleic acids encoding them are useful for  
 CC identifying modulators. The identified modulators are useful as  
 CC herbicides

XX Sequence 328 AA;

Query Match 8.7%; Score 86.5; DB 5; Length 328;  
 Best Local Similarity 25.8%; Pred. No. 1.4;  
 Matches 49; Conservative 18; Mismatches 86; Indels 37; Gaps 10;

QY 14 ITDASSLNGV-----DKLLSAEVEKMLVQKAPNEGIEVFGLLVLAARTSP 64  
 DB 105 LSDQILTVGCFAPAGIVGDKKAMEIINNAFVVSAGEND-----FILNYDIFSRLEY 159  
 QY 65 KVQRADSDVIFSN-----SFGERNVVTGDLKKVLDGCAPIRTFNKLTFTGRT 115  
 DB 160 PRISGVQDFILKRLNFVRELYSIGVRNVLV--GGLPFM--GCLPI--HMTAKFNIFRFC 214  
 QY 116 TEAYVDFCIAYKHKLPLNAAAEELGIPEDSYLAADFLGTCPELSELOOSRKMFAFWAL 175  
 DB 215 LEHNKRDVLTNEXKLQKLLPOLIEASLPF-SKFLVADVYN--PMMEMTONPSK-----YGF 266  
 QY 176 K-TEGGVNT 184  
 DB 267 KETKRGCCGT 276

RESULT 5  
 AAG39167  
 ID AAG39167 standard; protein, 295 AA.

XX AAG39167;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 48421.

XX Protein identification; signal transduction pathway; metabolic pathway;  
 KM hybridisation assay; genetic mapping; gene expression control; promoter;  
 KM termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 16-APR-1999; 99US-0128714P.

XX 19-APR-1999; 99US-0129845P.

XX 21-APR-1999; 99US-0130077P.

XX 23-APR-1999; 99US-0130510P.

XX 23-APR-1999; 99US-0130891P.

XX 28-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 30-APR-1999; 99US-0132407P.

XX 04-MAY-1999; 99US-0132484P.

PR	-MAY-1999	93UTS-0133485P
PR	06-MAY-1999	93UTS-0133486P
PR	06-MAY-1999	93UTS-0133487P
PR	07-MAY-1999	93UTS-0133488P
PR	11-MAY-1999	93UTS-0133258P
PR	14-MAY-1999	93UTS-0134218P
PR	14-MAY-1999	93UTS-0134219P
PR	14-MAY-1999	93UTS-0134221P
PR	14-MAY-1999	93UTS-0134370P
PR	15-MAY-1999	93UTS-0134768P
PR	18-MAY-1999	93UTS-0134941P
PR	20-MAY-1999	93UTS-0135124P
PR	21-MAY-1999	93UTS-0135353P
PR	24-MAY-1999	93UTS-0135629P
PR	25-MAY-1999	93UTS-0136021P
PR	27-MAY-1999	93UTS-0136392P
PR	28-MAY-1999	93UTS-0136782P
PR	01-JUN-1999	93UTS-0137222P
PR	03-JUN-1999	93UTS-0137528P
PR	04-JUN-1999	93UTS-0137502P
PR	07-JUN-1999	93UTS-0137724P
PR	08-JUN-1999	93UTS-0138094P
PR	10-JUN-1999	93UTS-0138540P
PR	10-JUN-1999	93UTS-0138847P
PR	14-JUN-1999	93UTS-0139111P
PR	16-JUN-1999	93UTS-0139452P
PR	16-JUN-1999	93UTS-0139453P
PR	17-JUN-1999	93UTS-0139492P
PR	18-JUN-1999	93UTS-0139554P
PR	18-JUN-1999	93UTS-0139645P
PR	18-JUN-1999	93UTS-0139456P
PR	18-JUN-1999	93UTS-0139457P
PR	18-JUN-1999	93UTS-0139458P
PR	18-JUN-1999	93UTS-0139459P
PR	18-JUN-1999	93UTS-0139660P
PR	18-JUN-1999	93UTS-0139661P
PR	18-JUN-1999	93UTS-0139462P
PR	18-JUN-1999	93UTS-0139463P
PR	18-JUN-1999	93UTS-0139750P
PR	18-JUN-1999	93UTS-0139763P
PR	21-JUN-1999	93UTS-0139817P
PR	22-JUN-1999	93UTS-0139899P
PR	22-JUN-1999	93UTS-0140053P
PR	23-JUN-1999	93UTS-0140355P
PR	23-JUN-1999	93UTS-0140585P
PR	24-JUN-1999	93UTS-0140695P
PR	24-JUN-1999	93UTS-0140823P
PR	29-JUN-1999	93UTS-0141081P
PR	30-JUN-1999	93UTS-0141887P
PR	01-JUL-1999	93UTS-0141842P
PR	01-JUL-1999	93UTS-0142154P
PR	02-JUL-1999	93UTS-0142055P
PR	06-JUL-1999	93UTS-0142390P
PR	08-JUL-1999	93UTS-0142803P
PR	09-JUL-1999	93UTS-0143220P
PR	12-JUL-1999	93UTS-0143277P
PR	13-JUL-1999	93UTS-0143542P
PR	14-JUL-1999	93UTS-0143624P
PR	15-JUL-1999	93UTS-0144005P
PR	16-JUL-1999	93UTS-0144085P
PR	16-JUL-1999	93UTS-0144086P
PR	19-JUL-1999	93UTS-0144325P
PR	19-JUL-1999	93UTS-0144331P
PR	19-JUL-1999	93UTS-0144332P
PR	19-JUL-1999	93UTS-0144333P
PR	19-JUL-1999	93UTS-0144334P
PR	19-JUL-1999	93UTS-0144335P
PR	20-JUL-1999	93UTS-0144332P
PR	20-JUL-1999	93UTS-0144384P
PR	21-JUL-1999	93UTS-0144814P
PR	21-JUL-1999	93UTS-0145086P
PR	22-JUL-1999	93UTS-0145088P
PR	22-JUL-1999	93UTS-0145085P

PR	22-JUL-1999	9905-0145087P
PR	22-JUL-1999	9905-0145089P
PR	22-JUL-1999	9905-0145122P
PR	23-JUL-1999	9905-0145145P
PR	23-JUL-1999	9905-0145218P
PR	23-JUL-1999	9905-0145224P
PR	26-JUL-1999	9905-0145276P
PR	27-JUL-1999	9905-0145513P
PR	27-JUL-1999	9905-0145518P
PR	27-JUL-1999	9905-0147302P
PR	05-AUG-1999	9905-0147182P
PR	05-AUG-1999	9905-0147260P
PR	05-AUG-1999	9905-0147303P
PR	06-AUG-1999	9905-0147461P
PR	06-AUG-1999	9905-0147483P
PR	09-AUG-1999	9905-0147953P
PR	10-AUG-1999	9905-0148111P
PR	11-AUG-1999	9905-0148319P
PR	12-AUG-1999	9905-0148341P
PR	13-AUG-1999	9905-0148556P
PR	15-AUG-1999	9905-0148664P
PR	16-AUG-1999	9905-0149368P
PR	17-AUG-1999	9905-0149175P
PR	18-AUG-1999	9905-0149426P
PR	20-AUG-1999	9905-0149723P
PR	20-AUG-1999	9905-0149929P
PR	23-AUG-1999	9905-0149902P
PR	23-AUG-1999	9905-0149930P
PR	25-AUG-1999	9905-0150566P
PR	26-AUG-1999	9905-0150884P
PR	27-AUG-1999	9905-0151065P
PR	27-AUG-1999	9905-0151066P
PR	27-AUG-1999	9905-0151080P
PR	30-AUG-1999	9905-0151303P
PR	31-AUG-1999	9905-0151438P
PR	01-SEP-1999	9905-0151930P
PR	01-SEP-1999	9905-0152363P
PR	10-SEP-1999	9905-0153070P
PR	13-SEP-1999	9905-0153758P
PR	15-SEP-1999	9905-0154018P
PR	16-SEP-1999	9905-0154039P
PR	20-SEP-1999	9905-0154477P
PR	22-SEP-1999	9905-0155139P
PR	23-SEP-1999	9905-0155466P
PR	24-SEP-1999	9905-0155659P
PR	26-SEP-1999	9905-0156458P
PR	29-SEP-1999	9905-0156596P
PR	04-OCT-1999	9905-0157117P
PR	05-OCT-1999	9905-0157513P
PR	06-OCT-1999	9905-0157865P
PR	07-OCT-1999	9905-0158029P
PR	08-OCT-1999	9905-0158233P
PR	12-OCT-1999	9905-0158369P
PR	13-OCT-1999	9905-0159293P
PR	13-OCT-1999	9905-0159294P
PR	13-OCT-1999	9905-0159295P
PR	14-OCT-1999	9905-0159328P
PR	14-OCT-1999	9905-0159330P
PR	14-OCT-1999	9905-0159331P
PR	14-OCT-1999	9905-0159637P
PR	14-OCT-1999	9905-0159638P
PR	18-OCT-1999	9905-0159584P
PR	21-OCT-1999	9905-0160741P
PR	21-OCT-1999	9905-0160767P
PR	21-OCT-1999	9905-0160768P

PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 22-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 26-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query March 8.7%; Score 86; DB 3; Length 295;  
Best Local Similarity 26.5%; Pred. No. 1.3;  
Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY 7 SNLSNLTVDASSINGV--DKLLISAEVERKMLYOKGAPNNGI---EVVFGILLYLAAR 60  
DB 72 SEQPNMFKSYARLKGIYDKKMEIINNAFVWVSAGPNDFINYYEIPSRRLYEYFISG 121  
QY 61 TTSPKVGARADSDVFFNSPFGRRNVVTEGDLKKVLDGCAPLTFRTNKLRTFGRTTEAYV 120  
DB 132 YQDPIILKLENFVRELKSLGVANLV--GGLPEW--GCLPI-HHTAKFRNIFRCLEHNN 186  
QY 121 DFCIAYRHKLPGQINAAAEELGIPADSYLADFLGTCTKLSLQSRMFMASMTALK-TEG 179  
DB 187 KDSLVNFKLQNLPLQTEASLPG-SKFLYADVNA--EMMEMIQVPSK-----YGFKETKR 238  
QY 180 GVWNT 184  
DB 239 GCCGT 243

## RESULT 6

AAG20735  
ID AAG20735 standard; protein, 295 AA.

XX AAG20735;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 23037.

XX Protein identification; signal transduction pathway; metabolic pathway;  
XX Hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.

XX Arabidopsis thaliana.

XX OS  
XX EP1033405-A2.XX PN  
XX 06-SEP-2000.XX FD  
XX 25-FEB-2000; 2000EP-00301439.XX PF  
XX 25-FEB-1999; 99US-0121825P.

XX PR 05-MAR-1999; 99US-0123180P.

XX PR 09-MAR-1999; 99US-0123548P.

XX PR 23-MAR-1999; 99US-0125788P.

XX PR 25-MAR-1999; 99US-0126264P.

XX PR 29-MAR-1999; 99US-0126785P.

XX PR 01-APR-1999; 99US-0127462P.

XX PR 06-APR-1999; 99US-0128234P.

XX PR 08-APR-1999; 99US-0128714P.

XX PR 16-APR-1999; 99US-0129845P.

XX PR 19-APR-1999; 99US-0130077P.

XX PR 21-APR-1999; 99US-0130449P.

XX PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 14-MAY-1999; 99US-0134372P.  
PR 18-MAY-1999; 99US-0134941P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136393P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139112P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 26-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142930P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142820P.  
PR 12-JUL-1999; 99US-0142877P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 18-JUL-1999; 99US-0144325P.  
PR 18-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.

PR 20-JUL-1999; 99US-014484P.  
 PR 21-JUL-1999; 99US-014481P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 25-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 23-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158028P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159931P.  
 PR 14-OCT-1999; 99US-0159637P.

PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 22-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161932P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 8.7%; Score 86; DB 3; Length 295;  
 Best Local Similarity 26.5%; Pred. No. 1.3;  
 Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY 7 SNLSNLYITDASSLNGV--DKGLSAEVEKMLYOKGAPNEGI---EIVFGLLLYLAAR 60  
 DB 72 SEQPMNFKSYIARLKGIVGDKKAMEIINNAPVAVSAGNDFILMYEIPSRRLPYPTSG 131  
 QY 61 TTSPPVOPADSDVYFNSNFGERNVVTEDGDKVLDGCAPIRTFTNKLRTGRTTEAYV 120  
 DB 132 YODFLIKLENVRLYSLGVRNVYV--GGLPPM--GGLPI--HMTAKFRNIFRCLBEHN 186  
 QY 121 DECIAYKKKLPQNLNAALVGIIPADSVYADPFLGTCPKLSLQSRKKFASMYALK--TEG 179  
 DB 187 KDSVLYNEKQLNLPQIBASLPG--SKFLYADVYN--PMMEMIQNPSK----YGFKEIKR 238  
 QY 180 GVVNT 184  
 DB 239 GCCGT 243

RESULT 7  
 AAG20734  
 ID AAG20734 standard; protein; 360 AA.  
 AC AAG20734;  
 XX  
 DT 17-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 23036.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KM hybridisation assay; genetic mapping; gene expression control; promoter;  
 XX termination sequence.  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-AA.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-00301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 04-APR-1999; 99US-0132407P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 06-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140822P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142800P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.

PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147499P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151067P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152362P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154799P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 04-OCT-1999; 99US-0156596P.  
 PR 05-OCT-1999; 99US-0157117P.  
 PR 06-OCT-1999; 99US-0157655P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158329P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.

PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 22-OCT-1999; 99US-0161404P.  
 PR 22-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161320P.  
 PR 28-OCT-1999; 99US-0161322P.  
 PR 28-OCT-1999; 99US-0161393P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 8.7%; Score 86; DB 3; Length 360;  
 Best Local Similarity 26.5%; Pred. No. 18;  
 Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY 7 SNLSNIVITDASSINCV--DKLLSAEVEKMLVQKGAENEGL---EYVFGLLLYALAAR 60  
 DB 137 SEOPNPFKXYIARLKGIVGDKKAMETINNAFVVSAGNDFLLNYEIPSRRLPYPTISG 196  
 QY 61 TTSFKVQADSVIFSNFSGERNVVTEGDLKKVLDGCAPLTRFTNKLRTGRTETAYV 120  
 DB 197 YQDFILKRLNFYRELISLVANVLY--GGLPEM--GGLPI--HMTAKFRNIFRCLBHN 251  
 QY 121 DFCIAKHLPLQNAALGIPAEBSYLAADPLGCPKLSELQCSRKMFAMYLK--TEG 179  
 DB 252 KQSVLNEKQNLPLQIEIASLPQ--SKFLYADVYN--PMMEMIQNPSK-----YGFKEKTKR 303  
 QY 180 GVNTT 184  
 DB 304 GCCGT 308

RESULT 8  
 AAG39166  
 ID AAG39166 standard; protein; 360 AA.

XX AAG39166;  
 XX  
 DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 48420.

KW Protein identification; signal transduction pathway; metabolic pathway;  
 hybridisation assay; genetic mapping; gene expression control; promoter;  
 termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127463P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0128845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 28-APR-1999; 99US-0130891P.  
 PR 30-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 04-MAY-1999; 99US-0132407P.  
 PR 05-MAY-1999; 99US-0132484P.  
 PR 06-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 07-MAY-1999; 99US-0132487P.  
 PR 11-MAY-1999; 99US-0132853P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134376P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138549P.  
 PR 10-JUN-1999; 99US-0138647P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139453P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140351P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142056P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145152P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147439P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148319P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149923P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157553P.  
 PR 06-OCT-1999; 99US-0157865P.

PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158346P.  
 PR 13-OCT-1999; 99US-0158293P.  
 PR 13-OCT-1999; 99US-0158294P.  
 PR 13-OCT-1999; 99US-0158295P.  
 PR 14-OCT-1999; 99US-0158329P.  
 PR 14-OCT-1999; 99US-0158330P.  
 PR 14-OCT-1999; 99US-0158331P.  
 PR 14-OCT-1999; 99US-0158337P.  
 PR 14-OCT-1999; 99US-0158638P.  
 PR 18-OCT-1999; 99US-0158584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 25-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 26-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 8.7%; Score 86; DB 3; Length 360;  
 Best Local Similarity 26.5%; Pred. No. 1.8;  
 Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY 7 SNLSNVIITDASSLNGV--DKLLSAEYKMKYQKGAPEGI---EYVFGLLLYALAA 60  
 DB 137 SEQPNMFSTYARKLGIYDCKAMEIINNAFVVSAGRNDFLNTYETPSRLLEFPFISG 196  
 QY 61 TTSPKVRADSDVIFSNSEFGERNVVTEGDLKVLGCAPLTRFTNKLRTFGRTTEAV 120  
 DB 197 YQDFIKRLNFREYSLGVNVLV--GGLPPM--GCLPI-HMTAKFRNIFRFLIEHKN 251  
 QY 121 DFCIAVKTLPOLNMAAEIGIPAEBSYLAADFLGTCPTKSEIQQRKMFASVATK-TEG 179  
 DB 252 KQSVLYNEKLNLLPQIEASLPG-SKFLYADVYN--PVMEMIQNPSK-----YGRKEIKR 303  
 QY 180 GVNT 184  
 DB 304 GCCGT 308  
 RESULT 9  
 ABB91962  
 ID ABB91962 standard; protein; 360 AA.  
 XX  
 AC ABB91962;  
 DT 31-MAY-2002 (first entry)  
 XX  
 DE Herbicidally active polypeptide SEQ ID NO 1173.  
 XX  
 KM Herbicidal; plant; agriculture; herbicide.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN W0200210210-A2.  
 XX  
 PD 07-FEB-2002.  
 XX  
 PF 28-AUG-2001; 2001MO-EP009892.  
 XX



PR 28-AUG-2001; 2001WO-EP009892.  
XX (FARB ) BAYER AG.  
XX Tiejfen K, Weidler M;  
PI WPI: 2002-269010/31.  
XX  
DR Identifying plant target proteins for herbicidally active compounds,  
PI comprising aligning and comparing nucleic acid or amino acid sequences  
PT from plant with nucleic acid or amino acid sequences from non-plant  
PT organisms.  
XX  
PS Claim 5; SEQ ID NO 1173; 261pp + Sequence listing; English.  
XX  
XX The invention relates to identifying target proteins (AB990790-AB994016)  
CC for herbicidally active compounds, comprising aligning and comparing  
CC nucleic acid or amino acid sequences from plant with nucleic acid or  
CC amino acid sequences from non-plant organisms using suitable search  
CC parameters, where plant sequences having an E-value greater by a factor  
CC of 3 than the E-value of most similar non-plant sequences are selected.  
CC The polypeptides or nucleic acids encoding them are useful for  
CC identifying modulators. The identified modulators are useful as  
CC herbicides  
XX  
SQ Sequence 360 AA;  
  
Query Match 8.7%; Score 86; DB 5; Length 360;  
Best Local Similarity 26.5%; Pred. No. 1.8; Mismatches 20; Gaps 9;  
Matches 49; Conservative 23; Indels 20; Gaps 9;  
  
QY 7 SNLSNLTVDASLNGV--DKKLLSAEYKMLVQKQADNEG1---EYVFEGLLYALAAAR 60  
DB 137 SEQPMFNSYIARLKGIVGDKKAMEIINNAFVWSAGPNDFLNYEIPSRLEVPPISG 196  
QY 61 TTSPTVQRDSVITSNSRGERNVVVTGDDKKVLDGAPLRFNKLRTGRTFTYAYV 120  
DB 197 YDDFTLKRLNFVRELYSIGVRNVLY--GGLPFM--GCLPI-HMTAKRNFRCLEHNN 251  
QY 121 DECIAYKHLPOLNAAAEIGIAPEDSYLAADFLGTCPKLSEIQSRKMFASMYALK-TEG 179  
DB 252 KDSVLYNEKQLQLPQIEASLPG-SKFLYADVYN--PMMEMIQNPCK-----YGFKETKR 303  
QY 180 GVVNT 184  
DB 304 GCCGT 308  
  
RESULT 10  
AAAG39165  
ID AAAG39165 standard; protein; 366 AA.  
XX  
AC AAAG39165;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 48419.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridization assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
XX  
XX Arabidopsis thaliana:  
OS  
XX  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
PF  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 28-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 22-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131495P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132483P.  
PR 11-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135153P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137528P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139464P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140053P.  
PR 23-JUN-1999; 99US-0140054P.  
PR 24-JUN-1999; 99US-0140055P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.

PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144332P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145132P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 26-JUL-1999; 99US-0145919P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151068P.  
PR 27-AUG-1999; 99US-0151068P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 13-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154035P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 26-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.

PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158234P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159328P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159337P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160747P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161924P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 8.7%; Score 86; DB 3; Length 366;  
Best Local Similarity 26.5%; Pred. No. 1.8; Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY 7 SNLSNVITTDASSINGV--DKKLSAEVEKKLVQKGFNEG1---EYVFGLLYYALNAR 60  
DB 143 SEQPMFKSYIARLKGIVGDKKAMEIINNAFVVSAGHNDFTLNYEIPSRLEYPFISG 202  
QY 61 TTSRKVQPADSDVITSNSFGERNVYVTEGDLKKVLDGAPLTFRNKJRTGRTTEAYV 120  
DB 203 YODFTUKRLNFVRLYSLGVANVAV--GGLPPM--GCLPI-HMTAKRNITFRCLFHHN 257  
QY 121 DPCIAVKHKLPLQNNAAEYKIPAEBSYLAADFGLGTCPKLSEIQSRKKNFASMYALK-TEG 179  
DB 258 KDSVLYNEKLVNLLPQIEASLPG-SKFLYADVYN--PMMEMIQNPSK-----YGFKEKTKR 309  
QY 180 GYVNT 184  
DB 310 GCCGT 314

RESULT 11  
AAG20733  
ID AAG20733 standard; protein; 366 AA.  
XX AAG20733;  
AC  
XX  
XX 17-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23035.  
XX  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
PD  
XX

Fri Apr 9 15:30:07 2004

us-09-613-486-15.rag

Page 12

PR	25-FEB-1999	9905-0121825P
PR	03-MAR-1999	9905-012180P
PR	03-MAR-1999	9905-0123548P
PR	23-MAR-1999	9905-0125768P
PR	23-MAR-1999	9905-0128264P
PR	25-MAR-1999	9905-0126752P
PR	01-APR-1999	9905-0126745P
PR	08-APR-1999	9905-0126234P
PR	16-APR-1999	9905-0128714P
PR	15-APR-1999	9905-0128845P
PR	19-APR-1999	9905-0130077P
PR	21-APR-1999	9905-0130449P
PR	23-APR-1999	9905-0130510P
PR	23-APR-1999	9905-0130891P
PR	23-APR-1999	9905-0131449P
PR	30-APR-1999	9905-0132048P
PR	30-APR-1999	9905-0132407P
PR	04-MAY-1999	9905-0132484P
PR	05-MAY-1999	9905-0132485P
PR	06-MAY-1999	9905-0132486P
PR	06-MAY-1999	9905-0133487P
PR	07-MAY-1999	9905-0133863P
PR	11-MAY-1999	9905-0134255P
PR	14-MAY-1999	9905-0134218P
PR	14-MAY-1999	9905-0134221P
PR	14-MAY-1999	9905-0134370P
PR	18-MAY-1999	9905-0134768P
PR	19-MAY-1999	9905-0134941P
PR	20-MAY-1999	9905-0135124P
PR	21-MAY-1999	9905-0135355P
PR	24-MAY-1999	9905-0135629P
PR	25-MAY-1999	9905-0136021P
PR	27-MAY-1999	9905-0136392P
PR	28-MAY-1999	9905-0136782P
PR	01-JUN-1999	9905-0137528P
PR	03-JUN-1999	9905-0137528P
PR	04-JUN-1999	9905-0137502P
PR	08-JUN-1999	9905-0138094P
PR	10-JUN-1999	9905-0138540P
PR	14-JUN-1999	9905-0138847P
PR	14-JUN-1999	9905-0139119P
PR	16-JUN-1999	9905-0139452P
PR	16-JUN-1999	9905-0139453P
PR	17-JUN-1999	9905-0139452P
PR	18-JUN-1999	9905-0139454P
PR	18-JUN-1999	9905-0139455P
PR	18-JUN-1999	9905-0139456P
PR	18-JUN-1999	9905-0139457P
PR	18-JUN-1999	9905-0139458P
PR	18-JUN-1999	9905-0139459P
PR	18-JUN-1999	9905-0139460P
PR	18-JUN-1999	9905-0139461P
PR	18-JUN-1999	9905-0139462P
PR	18-JUN-1999	9905-0139463P
PR	18-JUN-1999	9905-0139464P
PR	18-JUN-1999	9905-0139465P
PR	18-JUN-1999	9905-0139466P
PR	18-JUN-1999	9905-0139467P
PR	18-JUN-1999	9905-0139468P
PR	18-JUN-1999	9905-0139469P
PR	18-JUN-1999	9905-0139470P
PR	18-JUN-1999	9905-0139471P
PR	18-JUN-1999	9905-0139472P
PR	18-JUN-1999	9905-0139473P
PR	18-JUN-1999	9905-0139474P
PR	18-JUN-1999	9905-0139475P
PR	18-JUN-1999	9905-0139476P
PR	18-JUN-1999	9905-0139477P
PR	18-JUN-1999	9905-0139478P
PR	18-JUN-1999	9905-0139479P
PR	18-JUN-1999	9905-0139480P
PR	18-JUN-1999	9905-0139481P
PR	18-JUN-1999	9905-0139482P
PR	18-JUN-1999	9905-0139483P
PR	18-JUN-1999	9905-0139484P
PR	18-JUN-1999	9905-0139485P
PR	18-JUN-1999	9905-0139486P
PR	18-JUN-1999	9905-0139487P
PR	18-JUN-1999	9905-0139488P
PR	18-JUN-1999	9905-0139489P
PR	18-JUN-1999	9905-0139490P
PR	18-JUN-1999	9905-0139491P
PR	18-JUN-1999	9905-0139492P
PR	18-JUN-1999	9905-0139493P
PR	18-JUN-1999	9905-0139494P
PR	18-JUN-1999	9905-0139495P
PR	18-JUN-1999	9905-0139496P
PR	18-JUN-1999	9905-0139497P
PR	18-JUN-1999	9905-0139498P
PR	18-JUN-1999	9905-0139499P
PR	18-JUN-1999	9905-0139500P
PR	18-JUN-1999	9905-0139501P
PR	18-JUN-1999	9905-01

PR	03-JUL-1999;	99US-01429220P;
PR	12-JUL-1999;	99US-01428277P;
PR	13-JUL-1999;	99US-01435422P;
PR	14-JUL-1999;	99US-01433624P;
PR	15-JUL-1999;	99US-01440055P;
PR	16-JUL-1999;	99US-01440855P;
PR	16-JUL-1999;	99US-01440865P;
PR	19-JUL-1999;	99US-0144325P;
PR	19-JUL-1999;	99US-0144331P;
PR	19-JUL-1999;	99US-0144332P;
PR	19-JUL-1999;	99US-0144333P;
PR	19-JUL-1999;	99US-0144334P;
PR	19-JUL-1999;	99US-0144335P;
PR	19-JUL-1999;	99US-01443525P;
PR	20-JUL-1999;	99US-01446842P;
PR	20-JUL-1999;	99US-0144814P;
PR	21-JUL-1999;	99US-0145086P;
PR	21-JUL-1999;	99US-0145088P;
PR	22-JUL-1999;	99US-0145087P;
PR	22-JUL-1999;	99US-0145085P;
PR	22-JUL-1999;	99US-0145082P;
PR	22-JUL-1999;	99US-0145092P;
PR	23-JUL-1999;	99US-0145125P;
PR	23-JUL-1999;	99US-0145218P;
PR	23-JUL-1999;	99US-0145224P;
PR	26-JUL-1999;	99US-0145276P;
PR	26-JUL-1999;	99US-0145913P;
PR	27-JUL-1999;	99US-0145918P;
PR	27-JUL-1999;	99US-0145919P;
PR	28-JUL-1999;	99US-0145951P;
PR	02-AUG-1999;	99US-0146338P;
PR	02-AUG-1999;	99US-0146388P;
PR	02-AUG-1999;	99US-0146389P;
PR	03-AUG-1999;	99US-0147008P;
PR	04-AUG-1999;	99US-0147202P;
PR	04-AUG-1999;	99US-0147302P;
PR	05-AUG-1999;	99US-0147122P;
PR	05-AUG-1999;	99US-0147260P;
PR	06-AUG-1999;	99US-0147303P;
PR	06-AUG-1999;	99US-0147416P;
PR	09-AUG-1999;	99US-0147493P;
PR	09-AUG-1999;	99US-0147923P;
PR	09-AUG-1999;	99US-014811P;
PR	11-AUG-1999;	99US-0148311P;
PR	12-AUG-1999;	99US-0148341P;
PR	13-AUG-1999;	99US-0148552P;
PR	13-AUG-1999;	99US-0148684P;
PR	16-AUG-1999;	99US-0149336P;
PR	17-AUG-1999;	99US-0149175P;
PR	18-AUG-1999;	99US-0149426P;
PR	20-AUG-1999;	99US-0149722P;
PR	20-AUG-1999;	99US-0149723P;
PR	20-AUG-1999;	99US-0149929P;
PR	23-AUG-1999;	99US-0149902P;
PR	23-AUG-1999;	99US-0149930P;
PR	23-AUG-1999;	99US-0150566P;
PR	26-AUG-1999;	99US-0150884P;
PR	26-AUG-1999;	99US-0151065P;
PR	27-AUG-1999;	99US-0151066P;
PR	30-AUG-1999;	99US-0151086P;
PR	30-AUG-1999;	99US-0151303P;
PR	31-AUG-1999;	99US-0151438P;
PR	01-SEP-1999;	99US-0151930P;
PR	07-SEP-1999;	99US-0152363P;
PR	10-SEP-1999;	99US-0153070P;
PR	13-SEP-1999;	99US-0153758P;
PR	15-SEP-1999;	99US-0154018P;
PR	16-SEP-1999;	99US-0154032P;
PR	20-SEP-1999;	99US-0154377P;
PR	22-SEP-1999;	99US-0154513P;
PR	23-SEP-1999;	99US-0154566P;
PR	24-SEP-1999;	99US-0155482P;

PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 22-OCT-1999; 99US-0161404P.  
 PR 23-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161970P.  
 PR 28-OCT-1999; 99US-0161982P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query March 8.7%; Score 86; DB 3; Length 366;  
 Best Local Similarity 26.5%; Pred. No. 1.8;  
 Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY 7 SNISNLTTPASSINGV--DKKLISAEVEKMLVQKAPNNGI-----EVPFGILLVLAAR 60  
 DB 143 SEOPNMEKSYIARLKGIVGDKKMEIINNAFYVVSAGPNDFIINYYEIPSRRLYEYFISG 202  
 QY 61 TTSPKVGRAADSDVYFNSFGERNVVTGDLKKVLDGCAPLRTFTNKLRTFGRTFEAVY 120  
 DB 203 YQDFILKRLNFRRLYSLSGVNVLV--GGLPPV--GCLPI-IMTKFRNIFRQCEHNN 257  
 QY 121 DFCIAYGHKLPQLNAAALGIPADSYLADFLGTCKLSELQOSRMFMSTALK-TBG 179  
 DB 258 KDSVLYNEKIQNLPIQIEASLPG-SKFLYADVNN--FMEMIQIPSK-----YGFREYK 309  
 QY 180 GVAVT 184  
 DB 310 GCCGT 314

RESULT 12  
 AAB96072  
 ID AAB96072 standard; protein, 447 AA.

XX AC AAB96072;  
 XX DT 29-OCT-2001 (first entry)  
 XX DE Putative P. abyssi beta-lactamase.  
 XX KM Hyperthermophilic archaeon, hyperthermophilic protein.  
 XX OS Pyrococcus abyssi.  
 XX PN FR2792651-A1.

XX 27-OCT-2000.  
 PD 21-APR-1999; 99FR-0005034.  
 PF 21-APR-1999; 99FR-0005034.  
 PR 21-APR-1999; 99FR-0005034.  
 XX (CNRS) CNRS CENT NAT RECH SCI.  
 PA (IFRE)-IFREMER INST FR RECH EXPL MER.  
 XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O,  
 PI Querellou J, Weissenbach J, Saurin W, Helling R;  
 XX WPI, 2001-126236/14.  
 DR  
 XX  
 PT New nucleotide sequences isolated from Pyrococcus abyssi encode proteins  
 useful in industry.  
 PS  
 XX Claim 7, Page 693-694; 1657bp; French.  
 CC The present invention relates to the genomic sequence of Pyrococcus  
 CC abyssi (see AB86431 and AAH41223-7) and P. abyssi proteins. P. abyssi is  
 CC a hyperthermophilic archaeon, which is isolated from deep-sea  
 CC hydrothermal vents. The present sequence is one such P. abyssi protein.  
 CC The proteins of the present invention have various potential industrial  
 CC uses, since the proteins are stable at very high temperatures, some up to  
 CC 110 degrees centigrade. Note: This patent is in the same patent family as  
 CC WO200005062, which contains additional sequences as shown in AAB99132-  
 CC AAB99143, AAH75903-AAH75920 and AAG66436  
 XX

Query March 8.7%; Score 86; DB 4; Length 447;  
 Best Local Similarity 26.0%; Pred. No. 2.4;  
 Matches 40; Conservative 15; Mismatches 41; Indels 58; Gaps 8;

QY 26 KILSAEVRKMLVQKAPNNGIEVVPFGILLVLAARTSPKVRADSDVIFNSFGERNV- 84  
 DB 5 KLESPFVERKMERK-VPGISISII-----KQGVVYAKGFGYRNV 44  
 QY 85 -----VTEGDKKVLGCAPLRTFTN-XLRTFGRTFT 116  
 DB 45 ARLPSTPETIYGISITSTFTALIMKLVBEGLS--LDD--PYEKFNITLRPFGRVPT 100  
 QY 117 EAVYDFCIAYGHKLPQLNAAAL--GIPADSYL 148  
 DB 101 ---VHLLTHSSGIPSLGVAEAFIDGAVGDNWL 131

RESULT 13  
 AAH17788  
 ID AAH17788 standard; protein, 599 AA.

XX AC AAH17788;  
 XX DT 17-OCT-2003 (revised)  
 XX DT 19-AUG-1997 (first entry)  
 XX DE Phage abortive infection protein AbiE.  
 XX AbiE, abortive phage infection protein; phage resistance; pSRQ800;  
 XX lactic acid bacterium.  
 XX OS Lactococcus lactis subsp. lactis; isolate W1.  
 XX PN WO9720917-A2.  
 XX DT 12-JUN-1997.  
 XX DT 20-NOV-1996; 96WO-1B001385.  
 XX PR 01-DEC-1995; 95US-00565907.

PA (UNIL) QUEST INT BV.  
 PI Moineau S, Holler BJ, Vandenberg PA, Vedamuthu ER, Kondo JK;  
 DR WPI; 1997-319765/29.  
 N-PSDB; AAT68648.  
 XX  
 PT Isolated DNA encoding the Abie protein of *Lactococcus* - for protecting  
 PT strains used in production of fermented dairy products.  
 PS Claim 48; Page 30-32; 49pp; English.  
 XX  
 CC Abie (AAW17788) is a phage abortive infection protein that increases  
 CC resistance to phage. It is the expression product of an open reading  
 CC frame found in the 4.5 kb EcoRI fragment (AAT68648) of *Lactococcus lactis*  
 CC (L.L.) subsp. *lactis* w/ plasmid pSR0800. Abie acts at, or before, phage  
 CC replication and has no homology with known Abi proteins. It protects  
 CC against phages of 936 and P335 types, also against C2 when present on a  
 CC high copy number plasmid. Abie can be used to impart phage resistance to  
 CC bacteria, esp. L.1. that do not have natural resistance. The recombinant  
 CC bacteria can be used in the prodn. of fermented dairy products. (Updated  
 CC on 17-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 599 AA;  
 Query Match 8.6%; Score 85; DB 2; Length 599;  
 Best Local Similarity 24.3%; Pred. No. 4.7;  
 Matches 36; Conservative 26; Mismatches 46; Indels 40; Gaps 6;  
 QY 2 ELMSDSNL-----SLVITD-----ASSLNGVDR--KLSA 30  
 DB 261 EFLHBNLICBENMLINDNKTVDNPPYDKSKSDIFSFENITSTNDNMKIKESIN 320  
 QY 31 EWEKMLVQKAPNEG-IEVVFGLLYALAAARTSPKIQADSDVIFSNSEGERVVTTEG 89  
 DB 321 FIDYCVNEEHLGNKAIKCIPIVI-----TNTLKQKQVDTKNI-DNIFSKRMVTFNFN 372  
 QY 90 DLKRYLGGCAPLRTFTNKLRTFGRTFTT 117  
 DB 373 VFEKILDLSLKDSRLTNKLTFFENINE 400  
 XX  
 RESULT 14  
 AA97690  
 ID AA97690 standard; protein; 269 AA.  
 XX  
 AC AA97690;  
 XX  
 DT 06-AUG-2003 (revised)  
 DT 08-MAY-2001 (first entry)  
 XX  
 DE GRAY-5 coat protein.  
 XX  
 KW GRAY-5; grapevine leafroll virus; GRAY infection; GRAY coat protein;  
 KM GRAY HSP70 homologue protein; viral gene mapping;  
 KM plant disease resistance.  
 XX  
 OS Grapevine leafroll virus.  
 XX  
 PN WO200105957-A2.  
 XX  
 PD 25-JAN-2001.  
 XX  
 PF 19-JUL-2000; 2000WO-US019708.  
 XX  
 PR 19-JUL-1999; 99US-0144453P.  
 XX  
 PA (AGRI-) AGRITOPIC INC.  
 XX  
 PI Good XC, Monis J;  
 XX  
 DR WPI; 2001-147339/15.  
 DR N-PSDB; AAA91259.

XX  
 PT Novel grapevine leafroll virus polynucleotide useful as diagnostic and  
 PT probe, for viral gene mapping and for induced plant disease resistance.  
 XX  
 XX  
 PS Claim 1; Fig 1; 60pp; English.  
 XX  
 CC This sequence represents a grapevine leafroll virus (GRAY-5) protein  
 CC sequence of the invention. The DNA sequence can be used in an expression  
 CC construct. The construct is useful for providing resistance to GRAY  
 CC infection in a recombinant plant cell by transforming the plant cell with  
 CC it, where transcription of the polynucleotide sequence interferes with a  
 CC normal viral function such as movement, encapsidation or replication of  
 CC viral RNA. The polynucleotide sequence is expressed as an antisense  
 CC sequence and encodes a GRAY coat protein, preferably a defective GRAY  
 CC coat protein or a GRAY HSP70 homologue protein. The GRAY-5 DNA is  
 CC useful for the synthesis of GRAY, as diagnostic and probes, for viral  
 CC gene mapping and for induced plant disease resistance. It is also useful  
 CC to detect and quantitate expression of GRAY in plant tissue prior to use  
 CC in vegetative propagation, by detecting the presence of GRAY RNA.  
 CC (Updated on 06-AUG-2003 to correct OS field.)  
 XX  
 SQ Sequence 269 AA;  
 Query Match 8.5%; Score 84.5; DB 4; Length 269;  
 Best Local Similarity 24.4%; Pred. No. 1.7;  
 Matches 32; Conservative 21; Mismatches 57; Indels 21; Gaps 4;  
 QY 61 TTSPPVQADSDVIFSNSEGERVVTTEGDKVLDGAPLRTFTNKLRTFGRTFTTAYV 120  
 DB 127 STSPVSSSSNKTITGKIDGKRYVAHDIKTALDNGSGFENRTROGRATIAIV 185  
 QY 121 DFCIAVKKLPLNAAAEIGIPAE-----DSYLAADFGTCPKSEL 162  
 DB 186 QGISSGKLEV-NTKICASHGVFPNYPYSPDLHYDALFGYDASLAIE-LGKVAIVNKP 243  
 QY 163 QOSRCKMFASMY 173  
 DB 244 SNSNRATNLY 254  
 XX  
 RESULT 15  
 ABU26015  
 ID ABU26015 standard; protein; 422 AA.  
 XX  
 AC ABU26015;  
 XX  
 DT 19-JUN-2003 (first entry)  
 DT  
 XX  
 DE Protein encoded by prokaryotic essential gene #11542.  
 XX  
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX  
 OS *Corynebacterium diphtheriae*.  
 XX  
 PN WO200277183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002WO-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362599P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind W,  
 PI Wali D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX  
 DR WPI; 2003-029926/02.  
 DR N-PSDB; ACA29885.









capsid protein homolog p27 - citrus tristeza closterovirus CTV

C/Species: citrus tristeza closterovirus CTV

C/Date: 11-Oct-1994 #sequence\_revision 25-Apr-1997 #text\_change 11-Jan-2000

C/Accession: C49804

R/Pappu, H.R.; Karasev, A.V.; Anderson, E.J.; Pappu, S.S.; Hilt, M.E.; Febres, V.J.; Eck

S.M.; Dawson, W.O.; Lee, R.F.; Mblett, C.L.

Virology 199, 35-46, 1994

A/Title: Nucleotide sequence and organization of eight 3' open reading frames of the cit

A/Reference number: A49804; MUID:94160579; PMID:8116253

A/Accession: C49804

A/Status: preliminary

A/Molecule type: genomic RNA

A/Residues: 1-240 <PDP>

A/Cross-references: GB:U16304; GB:U02547; GB:L20760; NID:9806738; PIDN:AA059629.1; PID:9

A/Note: sequence extracted from NCBI backbone (NCBIN:144092, NCBI:P144099)

A/Note: severe quick decline isolate T36

C/Superfamily: SBV probable coat protein

Query Match 10.1%; Score 100.5; DB 2; Length 240;

Best Local Similarity 26.4%; Pred. No. 0.077;

Matches 57; Conservative 28; Mismatches 70; Indels 61; Gaps 13;

QY 2 EIMSDSNLNVIT-DASLNGVDKLLSAVEKMLVOKGAPNIEVFGLLYLAAR 60

DB 61 EKFTGEHLKVVWMDTFLPENTKTEDLLVLTMTQK-----RLYIT--- 104

QY 61 TTSPKRVADSDVIFSNFGRNVVTEGDLK-KVLDGCA--PLTRFT----NKLRTF 111

DB 105 STSKTKTFRDKGCI-----SVQGGRLRYLLKVVFPFLSKFTRETNPMLRKF 154

QY 112 GRTTEAVVDCIA-----YKHLTPOLNAAELGIPADSYLAADFL-CTCKLER--- 161

DB 155 ACTEEELHL-CMARLPDILENK-----RTTRAGTGLHLYGLSADPLSGLSYSEHER 207

QY 162 ---LQOSKRFMSYALKTEGGVNTFVSNLROLGR 194

DB 208 GILRASESMIA-----RQGYEATELNLNRDGLK 238

RESULT 7

hypothetical protein 5 - sugar beet yellows virus

C/Species: sugar beet yellows virus, SBV

C/Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 10-Dec-1999

C/Accession: S28714

R/Aganovsky, A.A.; Boyko, V.P.; Karasev, A.V.; Lunina, N.A.; Koonin, E.V.; Dolja, V.V.

J. Gen. Virol. 72, 15-23, 1991

A/Title: Nucleotide sequence of the 3'-terminal half of beet yellows closterovirus RNA

A/Reference number: S28710; MUID:9116305; PMID:1990061

A/Accession: S28714

A/Molecule type: DNA

A/Residues: 1-216 <AGR>

A/Cross-references: EMBL:X53462; NID:958678; PIDN:CAA31553.1; PID:958682

C/Superfamily: SBV probable coat protein

Query Match 10.0%; Score 99.5; DB 2; Length 216;

Best Local Similarity 25.0%; Pred. No. 0.083;

Matches 50; Conservative 27; Mismatches 94; Indels 29; Gaps 7;

QY 5 SDSNLSNIVITDASLNGVDKLLSAVEKMLVOKGAPNIEVFGLLYLAARTSP 64

DB 34 SEVPFNKLNKRETDGLGVIRERKSELV-----ITDEDFVHLAFLIRANITTSV 86

QY 65 KYQADSDVIFSNFGRNVVTEGDLKVDGCAPLTRFNTKRTGRTFTEAVVDFCI 124

DB 87 KVVNYGA---YEYITGKKFLVKNAMVPELIKCKKKNKPNVPTCATPEADY----I 139

QY 125 AYKHLPOL---NAAELGIPADSYLAADFL--GTC---PKLSELQOSKRFMSMYA 174

DB 140 VIASLPLFLNRTIGKGIPIGDFIGADPLTATSVCLNDHEKAVIAYLSRAADRAYS 199

QY 175 LKTEGGVNTFVSNLROLGR 194

DB 200 SSVDDKIV-----SLFDLGR 214

RESULT 8

A/Accession: A33988

A/Species: Schizosaccharomyces pombe

C/Date: 23-Mar-1990 #sequence\_revision 23-Mar-1990 #text\_change 31-Mar-2000

C/Accession: A33988; A33539; T39809

R/Young, D.; Riggs, M.; Field, T.; Vojtek, A.; Broek, D.; Wiegler, M.

Proc. Natl. Acad. Sci. U.S.A. 86, 7989-7993, 1989

A/Title: The adenylate cyclase gene from Schizosaccharomyces pombe.

A/Reference number: A33988; MUID:90046723; PMID:2682634

A/Accession: A33988

A/Molecule type: DNA

A/Residues: 1-1692 <YOU>

A/Cross-references: GB:M26699; NID:G17338; PIDN:AA35284.1; PID:G17339

R/Yamawaki-Kataoka, Y.; Tamaoki, T.; Choe, H.R.; Tanaka, H.; Kataoka, T.

Proc. Natl. Acad. Sci. U.S.A. 86, 5693-5697, 1989

A/Title: Adenylate cyclases in yeast: a comparison of the genes from Schizosaccharomyces

A/Reference number: A33539; MUID:89345533; PMID:2669944

A/Accession: A33539

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1692 <YAM>

A/Cross-references: GB:M24942; NID:G17338; PIDN:AA35301.1; PID:G17339

A/Note: the authors translated the codon TGC for residue 626 as Ser, and GCC for residue

R/Lyne, M.; Rajandream, M.A.; Bartell, B.G.; Lucas, M.; Galliard, C.

submitted to the EMBL Data Library, June 1998

A/Reference number: Z21881

A/Accession: T39809

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1692 <LYN>

A/Cross-references: EMBL:AL023859; PIDN:CAA19571.1; GSPDB:GN00067; SPDB:SPBC19C7.03

A/Experimental source: strain 972h-, cosmid c19C7

C/Genetics:

A/Map position: 2

C/Superfamily: yeast adenylate cyclase; leucine-rich alpha-2-glycoprotein repeat homolog

C/Keywords: phosphorus-oxygen lyase

F/1328-1413/Domain: yeast adenylate cyclase catalytic domain homology <YACC>

Query Match 9.3%; Score 92; DB 2; Length 1692;

Best Local Similarity 24.7%; Pred. No. 5.9;

Matches 42; Conservative 26; Mismatches 66; Indels 34; Gaps 7;

QY 27 LLSAEVEKMLVOKGAPNIEVFGLLYLAARTSPKVRADSDVIFSNFGRNVV 86

DB 339 LIQFTEIRILL-----PHEQPCIFERLLSLFCKKTS-----DEINEEDNYSVARLVF 388

QY 87 TEGDLKVVLDGCAPLTRFNTKRT---FGRFTEAVVDFCIAYGKLPOLNAAELGIP 142

DB 389 TMDI-----GADVLRKRSKKTITANLDSNLEVIPIKYIPYHLLISNLSNLSLD 443

QY 143 ADSVYLAADFLGTCPLSELQOSKRFMSYALKTEGGVNTFVSNLROL 192

DB 444 -----LPIDFMERCVALKRLDISNN-----LRSPRG--KFTVALROL 478

RESULT 9

AD3246

P-450 monooxygenase vtrH [imported] - Agrobacterium tumefaciens (strain C58, Dupont) pl.

C/Species: Agrobacterium tumefaciens

C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002

C/Accession: AD3246

R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Moks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.

erage, G.; Giller, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McClell,

Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193  
 A:Accession: AD3246  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-419 <KIR>  
 A:Cross-references: GB:AE008690; PIDN:AAL46386.1; PID:g17744179; GSPDB:GN00189  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: vtrH  
 A:Genome: plasmid  
 C:Superfamily: Agrobacterium plasmid cytochrome P450 pinF; cytochrome P450 homology  
 C:Keywords: heme; iron; metalloprotein  
 F:366/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 9.1%; Score 90; DB 2; Length 419;  
 Best Local Similarity 25.1%; Pred. No. 1.5; Mismatches 71; Indels 58; Gaps 11;  
 Matches 52; Conservative 26

QY 20 LNVGDKLLSAEYKMLVOKGAPNKGIE--VPEGLLYALAASTSPKQVQADSDVIFSN 77  
 Db 67 LIGDPRTRQETELML-----NRGVRGAVFDLIRYSMLFNGEYHVRSA---FAK 117  
 QY 78 SFGRRNVVTEGDLKAVLDGAP-LTRFTNKLRTFGRTFEAVDPCIAVKHLPOLNA 136  
 Db 118 TFAFR-----MIDALRPEITKLTLEHMDVPRVD--FDFAEMTASKLPAITTA 164  
 QY 137 AELGIPAD-----SYLADFLGTCPKLSHQ-----QSRKM-- 168  
 Db 165 SVLGLPQDAFFTRIVYVNSRCLSPMGEDDPFEIBASAVELQDYVRAVAVDRSRRI 224  
 QY 169 -FASMY--ALKTGGVNTVPVSNLROJ 192  
 Db 225 DFLSCYLKAVREGLT--SPLEIMQL 249

RESULT 10  
 phosphoenolpyruvate carboxykinase pckA - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 27-Oct-2003  
 C:Accession: F69673  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berber, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chik, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 350, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, A.; Hore, J.; Harwood, C.R.; Hensut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.; Koester, P.; Koningsstein, G.; Krogsh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, T.M.; Portecelle, Rieger, M.; Rivola, C.; Rooha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeder, R.; Scoffone, F.; Scoffone, P.; Tognoni, A.; Tosato, K.; Uchiyama, T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumestein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis. A:Reference number: A69580; MUID:98044033; PMID:9384377  
 A:Accession: F69673  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-557 <KUN>  
 A:Cross-references: GB:Z99119; GB:AL009126; NID:92635411; PIDN:CB15034.1; PID:g2635540  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: pckA  
 C:Superfamily: phosphoenolpyruvate carboxykinase [ATP]

Query Match 8.8%; Score 87; DB 2; Length 527;  
 Best Local Similarity 26.9%; Pred. No. 3.9; Mismatches 47; Conservative 24; Indels 34; Gaps 9;  
 Matches 47; Conservative 24

QY 20 LNVGDKLLSAEYKMLV--QKGAPEGIEVFGILLALAASTSPKQVQADSDVIFS 76

Db 234 LSGTKTLLSADNRKLLGDEHGWSDTVGFNIEG-GCYAKCIHLSBEK---EPQIFNA 288  
 QY 77 NSFGE--RNVVTEGDLKAVLDGAP-LTRFTNKLRTFGRTFEAVDPCIAVK-HKLPOL 133  
 Db 289 IRFGSVLENVYVDD-----TRANVDDSPYENTRA-----AYPIHMINNI 330

QY 134 NAAELGIPADSYLADPLGTCPKLSHQSRM--FASMYALK--TEGGVNV 183  
 Db 331 VTPSWAGHPSAIVPLTADAFGLVPLISKLTKEQVYHFLSGYTSKLAGTERGVT 385

RESULT 11  
 B84722  
 probable GDSL-motif lipase/hydrolase [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
 C:Accession: B84722  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; Euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana. A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: B84722  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-328 <STO>  
 A:Cross-references: GB:AE002093; NID:g4582450; PIDN:AAD24834.1; GSPDB:GN00139  
 A:Gene: At2g31550  
 A:Map position: 2  
 C:Superfamily: myrosinase-associated protein MyAP

Query Match 8.7%; Score 86.5; DB 2; Length 328;  
 Best Local Similarity 25.8%; Pred. No. 2.3; Mismatches 49; Conservative 18; Indels 37; Gaps 10;  
 Matches 49; Conservative 18

QY 14 ITDASLNGV-----DKKLSAEYKMLVOKGAPNKGIEVFGILLALAASTSP 64  
 Db 105 LBDQDILTVCPASAGIVGDKKAMEIINNAFVVSAGPND-----FILNYVDIBSRLEY 159  
 QY 65 KYQVADSDVIFSN-----SFGRRNVVTEGDLKAVLDGAP-LTRFTNKLRTFGRTF 115  
 Db 160 PFIQGVDFILRLNFVRELTVLGRVNV--GGLPVP--GCLPI-HMTAKFRIRFFC 214  
 QY 116 TEAVVDFCAVYGHKLPQLNAAELGIPADSYLADPLGTCPKLSHQSRMFAKAYAL 175  
 Db 215 LEHHNNDVLYNEKQKLPQIEASLPG-SKFLVADVNV--PMEMIONPSK-----YGF 266  
 QY 176 K-TEGGVNT 184  
 Db 267 KETKRCGCGT 276

RESULT 12  
 A84722  
 probable GDSL-motif lipase/hydrolase [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
 C:Accession: A84722  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; Euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana. A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: A84722  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-360 <STO>  
 A:Cross-references: GB:AE002093; NID:g4582449; PIDN:AAD24833.1; GSPDB:GN00139  
 A:Gene: At2g31540

A:Map position: 2  
C:Superfamily: myrosinase-associated protein MYAP

Query Match 8.7%; Score 86; DB 2; Length 360;  
Best Local Similarity 26.5%; Pred. No. 2.9;  
Matches 49; Conservative 23; Mismatches 93; Indels 20; Gaps 9;

QY 7 SNLSNLTVDASSINGV--DKLLSAVEKMLVOKGAPNEGIV----EVLGILLYALAA 60  
DB 137 SEQPMFKASYIARKIGIVGDKAMEIINNAFVVVAGPNDFILNYEIPSRRLLEPPTISG 196  
QY 61 TTSPKVRADSDVIFNSFGERNVVTGDLKVLDCAPLTFPTNKLRTFGRTFEAVY 120  
DB 197 YQDPIKLKLENFVLEISLVGRNVLV--GGLPRM--GCLPI--HMTAKFRNIFPFCLEHNN 251  
QY 121 DFCIAYKHLFQNLAAALGLIPADSDYLAADFLGTCRKLSELOQRKMFASVALK--TEG 179  
DB 252 KDSVLYMEKQLQNLPLQLEASLPQ--SKFLYADVYN--FMEMIQNPSK----YGFKETKR 303  
QY 180 GVAVNT 184  
DB 304 GCCGT 308

RESULT 13

hypothetical protein T12A2.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T34352  
R:Latreille, P.  
submitted to the EMBL Data Library, March 1996  
A:Description: The sequence of C. elegans cosmid T12A2.  
A:Reference number: Z21511  
A:Accession: T34352  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-445 <LAT>  
A:Cross-references: EMBL:U13019; PIDN:AAC24449.1; GSPDB:GN00021; CESP:T12A2.1  
A:Experimental source: strain Bristol N2; clone T12A2  
C:Genetics:  
A:Gene: CESP:T12A2.1  
A:Map position: 3  
A:Introns: 33/3; 117/3; 146/3; 202/3; 250/2; 297/3; 371/1; 407/2

Query Match 8.7%; Score 86; DB 2; Length 445;  
Best Local Similarity 24.7%; Pred. No. 3.9;  
Matches 54; Conservative 41; Mismatches 80; Indels 44; Gaps 12;

QY 1 MEIWSDSLNLVITDAS-----SINGVDKLLSAVEKMLVOKGAPNEGIV---- 48  
DB 32 IKILNDSNSNLTVDITNGKFSYIGNLNGAEKKGEGEVENLKIIDSNGGIAIPGV 91  
QY 49 -----VF--GILLYALA---ARTTSPKVRADSDVIFNSFGERNVVTGDLKVLDCGA 99  
DB 92 GHSNHFVFGDRHREHMLAGATVNEVQAGGIGFTTN---KTRKASEQDLRKDFEELA 148  
QY 100 PLTRFTNKLRTGRTFGRTFEAVYDFCIAVYKHLFQNL--AAALGLIPADSDYLAADFLG--T 155  
DB 149 -----KMLRSGITTLLEKSGYGLNVDAMKMLVLTENPNPLE--VSATFCGAHA 199  
QY 156 CPKLS--ELQOSR-----KMPASVYALKTEGGVNTPEVNL 189  
DB 200 VPKSGTEYQTMICEELIPKIEDERKNGGLKN--VENI 236

RESULT 14

g75201  
probable beta-lactamase (EC 3.5.2.6) PAB0087 - Pyrococcus abyssi (strain Orsay)  
N:Alternate names: penicillin-binding protein homolog  
C:Species: Pyrococcus abyssi  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 05-Nov-1999  
C:Accession: G75201

R:Anonymous, Genoscope  
submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure  
A:Reference number: A75001  
A:Accession: G75201  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-447 <KAN>  
A:Cross-references: GB:A1248283; GB:A1096836; NID:95457433; PIDN:CA849062.1; PID:e1514951  
A:Experimental source: strain Orsay  
C:Genetics:  
A:Gene: bpd; PAB0087  
C:Keywords: hydrolase

Query Match 8.7%; Score 86; DB 2; Length 447;  
Best Local Similarity 26.0%; Pred. No. 3.9;  
Matches 40; Conservative 15; Mismatches 41; Indels 58; Gaps 8;

QY 26 KLLSAVEKMLVOKGAPNEGIVFGLLYALAAARTTSPKVRADSDVIFNSFGERNV-- 84  
DB 5 KLSFIVEKMAERK--VFGISISII-----KQGDVYVAKGFGRVAVE 44  
QY 85 -----VTEGDLKVLDCAPLTFPTN--KLRTFGRTFT 116  
DB 45 ARLPSTPRTTYIGISITGSFTALAIKLYEBSGLS--LDD--PVEKFNLIKLRPPGEPT 100  
QY 117 EAVYDFCIAVYKHLFQNLAAAL--GIPADSYL 148  
DB 101 ---VHLLTHSSGISPLGYABAFIDGMVGGDWML 131

RESULT 15

hypothetical protein b2627 - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 01-Mar-2002  
C:Accession: T08639; E65041  
R:Punkett, G.  
submitted to the EMBL Data Library, September 1995  
A:Reference number: Z16465  
A:Accession: T08639  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-729 <PLU>  
A:Cross-references: EMBL:U36840; NID:g1033110; PID:g1033122  
A:Experimental source: strain K12, substrain MG1655  
R:Blattner, F.R.; Punkett, G.; Bloch, C.A.; Ferna, N.T.; Burland, V.; Riley, M.; Co.  
A:Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; NCID:97426617; PMID:9278503  
A:Accession: E65041  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-729 <BLAT>  
A:Cross-references: GB:A6000348; GB:U00096; NID:g1788975; PIDN:AAC75675.1; PID:g1788980;  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:

A:Map position: 57 min  
Query Match 8.6%; Score 85.5; DB 2; Length 729;  
Best Local Similarity 22.4%; Pred. No. 8.1;  
Matches 41; Conservative 31; Mismatches 64; Indels 47; Gaps 8;

QY 2 ELMSDSLNLVIT--TDASSINGVDKLLSAVEKMLVOKGAPNEGIVFGLLYALAA 59  
DB 520 EAMDNDNFSSLVFHHMTATNTFLTKTIAKILAVLPHTFSRNGVVKPTDVIATLAGYLSA 579  
QY 60 RTTSPKVRADSDVITSN---SFGERNV--VTEGDLKVLDCAPLTFPTNKLRTFGRT 114  
DB 580 ESYSEYLRK---NQIDYARQWISGEKRTLSIALNNDLKI-----TN---TFGTY 623  
QY 115 FTAA-----YDFCIAVY--HKLIPQNLAAALGLIPADSYLA 150

Fri Apr 9 15:30:08 2004

us-09-613-486-15.rpr

Page 6

Db	624	LPKVLSTMEDVVKHAAVKRGIRSKVDYTHVKLA	PFHLEIPPGVNALEIGIPQTLRLV	683
Qy	151	DFL	153	
Db	684	DLT	686	

Search completed: April 5, 2004, 17:54:42  
Job time : 22 secs







[3]  
 RN SEQUENCE OF 1-165 FROM N.A.  
 RP STRAIN=168 / PY79;  
 RC MEDLINE=96345628; PubMed=8755891;  
 RX Youm R., Perkins J.B., Howitt C.L., Pero J.,  
 RA "Cloning and characterization of the mcs gene encoding S-  
 RT adenosulphomethionine synthetase from *Bacillus subtilis*."  
 RL J. Bacteriol. 178:4604-4610(1996).  
 CC -1- CATALYTIC ACTIVITY: ATP + oxaloacetate = ADP + phosphoenolpyruvate  
 CC + CO(2).  
 CC -1- PATHWAY: Rate-limiting gluconeogenic enzyme.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase (ATP)  
 CC family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF008220; AAC00377.1; -  
 CC EMBL; Z99119; CAB15034.1; -  
 CC EMBL; U52812; AAB17065.1; -  
 CC PIR; F69673; P69673.  
 CC HSSP; P22259; IAYL.  
 CC Subtilist; Bg1841; pcka.  
 CC HAVAP; MF\_00453; -; 1; PEPCK\_ATP.  
 CC InterPro; IPR001272; PEPCK\_ATP.  
 CC Pfam; PF01293; PEPCK\_ATP; 1.  
 CC ProDom; PD004723; PEPCK\_ATP; 1.  
 CC TIGRFAMs; TIGR00224; pcka; 1.  
 CC DR PROSITE; PS00532; PEPCK\_ATP; 1.  
 CC Gluconeogenesis; Lyase; Decarboxylase; ATP-binding; Complete proteome.  
 CC NP\_BIND; 233; 240; ATP (By similarity).  
 CC CONFLICT; 10; 10; L -> S (IN REF. 3).  
 CC SEQUENCE 527 AA; 58300 MW; E51EEC802D1E666 CRC64;  
 Query Match 8.8%; Score 87; DB 1; Length 527;  
 Best Local Similarity 26.9%; Pred. No. 2.6;  
 Matches 47; Conservative 24; Mismatches 70; Indels 34; Gaps 9;  
 QY 20 LNVGVKLLSAEVEKMLV---QKGAPEGIEVVGGLIYLAATAATTSBKVRADSDVIFS 76  
 DB 234 LSGTGKTLISADADKLTIGDEHGWSDYGVNIEG-GCYAKCIHISEK---EPQIFNA 288  
 QY 77 NSPGE--RNVVYTEGDLKKVLDGCAPLFRFNKLTFRGTETEAIVDFCIAYK-HKLPOL 133  
 DB 289 IRFGVLENVYVDE-----TREANVYDSSFTEENTRA-----AYPIHMINNI 330  
 QY 134 NAAAEGLIPADDSYIADFLGTCPKLSLQOSRKM--FASNYALK---TEGGVYN 183  
 DB 331 VTPSWAGHPSAIVFLTADAFGLTLPISKLTKEQVYHPLSGYTSKLACTGERTGVS 385  
 RESULT 5  
 YFUK\_ECOLI  
 ID YFUK\_ECOLI STANDARD; PRT; 729 AA.  
 AC P52126;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein yfjK.  
 GN YFUK OR B2627  
 OS *Escherichia coli*.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; *Escherichia*.  
 OC NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MC1655;

RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of *Escherichia coli* K-12."  
 RL Science 277:1453-1474(1997).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U36840; AAA79796.1; -  
 CC EMBL; AE000348; AAC75675.1; -  
 CC PIR; T08639; T08639.  
 CC EcoGene; EG13197; yfjK.  
 CC InterPro; IPR001410; DEAD.  
 CC InterPro; IPR001650; Helicase\_C.  
 CC Pfam; PF00271; helicase\_C; 1.  
 CC SMART; SM00487; DEXDC; 1.  
 CC SMART; SM00490; HELIC\_C; 1.  
 CC Hypothetical protein; Complete proteome.  
 CC SEQUENCE 729 AA; 83061 MW; C865BCD1B36PFAF CRC64;  
 Query Match 8.6%; Score 85.5; DB 1; Length 729;  
 Best Local Similarity 22.4%; Pred. No. 5.2;  
 Matches 41; Conservative 31; Mismatches 64; Indels 47; Gaps 8;  
 QY 2 ELMDSNLSNLYT--TDASSLNGVKKLLSAEVEKMLVQKGAPEGIEVVGGLIYLA 59  
 DB 520 EMMDDNEFSLVFPATNFKTKTAKIARLVPTFSRNGVYKPTVMIAKLAGIYSTA 579  
 QY 60 RTSPKVPQPADSDVIFSN---SFGERNV--VTEGDLKKVLDGCAPLFRFNKLTFRGT 114  
 DB 580 EGVSEYLRK---NQDIYARQWISGEKRTLSIANNNDLKI-----TN--TFGYT 623  
 QY 115 FTBA-----YDFCIAYK--HKLPOLNAAAEGLIPADDSYI 150  
 DB 624 LPKVSLMEDVYKHAVKRGRSKYDTHVKLAFESFHPGVNLEBIGIPITLHPLV 683  
 QY 151 DFL 153  
 DB 684 DLI 686  
 RESULT 6  
 PPCK\_RHOPA  
 ID PPCK\_RHOPA STANDARD; PRT; 537 AA.  
 AC O92NH4;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 26-FEB-2003 (Rel. 41, Last annotation update)  
 DE Phosphoenolpyruvate carboxykinase [ATP] (EC 4.1.1.49) (PEP  
 DE carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK).  
 GN PCKA.  
 OS *Rhodospseudomonas palustris*.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Bradyrhizobiaceae; *Rhodospseudomonas*.  
 OC NCBI\_TaxID=1076;  
 RN [1]  
 RP SEQUENCE FROM N.A., SEQUENCE OF 1-12, AND DEVELOPMENTAL STAGE.  
 RC STRAIN=7;  
 RC MEDLINE=99235744; PubMed=10217755;  
 RA Inui M., Nakata K., Roh J.H., Zahn K., Yukawa H.;  
 RA "Molecular and functional characterization of the *Rhodospseudomonas*  
 RT *palustris* no. 7 phosphoenolpyruvate carboxykinase gene."  
 RL J. Bacteriol. 181:2685-2696(1999).  
 CC -1- CATALYTIC ACTIVITY: ATP + oxaloacetate = ADP + phosphoenolpyruvate  
 CC + CO(2).



```

CC -1- PATHWAY: Rate-limiting gluconeogenic enzyme.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGH LEVELS DURING LOG PHASE
CC WITH 10-20 FOLD REDUCTION AT ONSET OF STATIONARY PHASE.
CC -1- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase (ATP)
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB015618; BAA34956.1; -.
DR HSSP; P22259; 1A02.
DR HAMAP; MF_00453; -. 1.
DR InterPro; IPR001272; PEPCK_ATP.
DR Pfam; PF01293; PEPCK_ATP; 1.
DR ProDom; PD004723; PEPCK_ATP; 1.
DR TIGRFAMs; TIGR00224; pckA; 1.
DR PROSITE; PS00532; PEPCK_ATP; FALSE NEG.
DR Glucomeogenesis; Lyase; Decarboxylase; ATP-binding.
KW NP BIND 236 243 ATP (BY SIMILARITY).
SQ SEQUENCE 537 AA; 58773 MW; 98F0E1CED1C9EAB CRC64;

Query Match 8.5%; Score 84; DB 1; Length 537;
Best Local Similarity 26.7%; Pred. No. 4.9;
Matches 48; Conservative 21; Mismatches 73; Indels 38; Gaps 10;

QY 20 LNVGDKLLSAEVENMLVQKAPNEGIEVFGLD-LYALAAFTTS--PRVQPADSDVI 74
DB 237 LSGGKTLTSDPNRTILGDDEHGMKQGVNPEFGCCAKIKLSANPEPIYVA----- 291
QY 75 FSNSEGE--RNVTYEGDLKYLVDGACAPLRFYTKLRTFGRTTEAYVDFCIANK-HKLP 131
DB 292 -STRFGAVLENNVLCGIDIRKPDFDGSK-----TENTRS-----AYPLESIP 332
QY 132 QMAAAHEGIPADESYLAADPLGCPKSEIQGSKX--FASWYALK---TEGGVVA-TP 185
DB 333 NASLTGRAGQPNVVMALADAFGWMPJAKLTPQAOMHFLSGTAKYAGIERGVTEPT 392

RESULT 7
TILD BUCAI STANDARD; PRT; 483 AA.
ID TILD BUCAI
AC P57478
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DR TILD protein homolog.
GN TILD OR Buj98.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS symbiotic bacterium).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=118093;
RN SEQUENCE FROM N.A.
RP STRAIN=Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RA "Genome sequence of the endocellular bacterial symbiont of aphids
RT Buchnera sp. APS."
RL Nature 407:81-86(2000).
CC -1- SIMILARITY: Belongs to the tild/pmba family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP001119; BAB3101.1; -.
DR InterPro; IPR002510; PcpC_dase_U62.
DR Pfam; PF01523; Pmba_TILD; 1.
KW Complete proteome.
SQ SEQUENCE 483 AA; 52977 MW; A05CE98518720EBA CRC64;

Query Match 8.4%; Score 83.5; DB 1; Length 483;
Best Local Similarity 20.5%; Pred. No. 4.8;
Matches 32; Conservative 26; Mismatches 53; Indels 45; Gaps 5;

QY 65 KYGRADSDVIFNSFGERNVYTEGDLKKVD-----GCAPLTR 103
DB 305 KNGRGSLSIDDEGTGQGNILKNGILKKYQDKNAALMVKNGNGRRESYCLPMWR 364
QY 104 FYNKRTFGRTTEAYVDFCIAYKHLPOLNAAELGIPADSYLAADPLGCPULSLQ 163
DB 365 MTN-----TYMLSGSKLDDIISKVDYGI-----YAVNFG--GQVDIT 401
QY 164 QSRKWPASWYALKTEGGVNVFVSNLROLGR-REYV 198
DB 402 SGRFVSTSEAYLIKNGKIVIPDKNTTLIGSLEVM 437

RESULT 8
ARGI LEPIN STANDARD; PRT; 385 AA.
ID ARG LEPIN
AC Q8EY78;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Arginine biosynthesis bifunctional protein argi [includes: Glutamate
DE N-acetyltransferase (EC 2.3.1.35) (Ornithine acetyltransferase)
DE (Ornithine transferase) (OATase); Amino-acid acetyltransferase
DE (EC 2.3.1.1) (N-acetylglutamate synthase) (AGS)] [contains: Arginine
DE biosynthesis bifunctional protein argi alpha chain; Arginine
DE biosynthesis bifunctional protein argi beta chain].
GN ARGJ OR LA4105.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN SEQUENCE FROM N.A.
RP STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi W.-H., Chen Z.,
RA Xu J.-G., Zhao G.-P.;
RT "Unique physiological and pathogenic features of Leptospira
RT interorgans revealed by whole-genome sequencing."
RL Nature 422:888-893(2003).
CC -1- FUNCTION: Catalyzes two activities which are involved in the
CC cyclic version of arginine biosynthesis: the synthesis of
CC acetylglutamate from glutamate and acetyl-CoA, and of ornithine by
CC transacetylation between acetylornithine and glutamate (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: N(2)-acetyl-L-ornithine + L-glutamate = L-
CC ornithine + N-acetyl-L-glutamate.
CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + L-glutamate = CoA + N-acetyl-L-
CC glutamate.
CC -1- PATHWAY: Arginine biosynthesis; first step.
CC -1- SUBUNIT: Heterotetramer of two alpha and two beta chains (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- MISCELLANEOUS: Some bacteria possess a monofunctional argi, i.e.,
CC capable of catalyzing only the fifth step of the arginine

```

```

CC biosynthetic pathway.
CC -1- SIMILARITY: Belongs to the argJ family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE011564; AANS1303.1; -.
CC HAMAP: MF_01106; -; 1.
CC InterPro: IPR002813; ArgJ.
CC Pfam: PF01960; ArgJ; 1.
CC ProDom: PD004193; ArgJ; 1.
CC TIGRFams: TIGR00120; ArgJ; 1.
CC Arginine biosynthes; Multifunctional enzyme; Transferase;
CC Acyltransferase; Complete proteome.
CC CHAIN 1 178
CC ARGININE BIOSYNTHESIS BIFUNCTIONAL
CC PROTEIN ARGJ ALPHA CHAIN (BY SIMILARITY).
CC ARGININE BIOSYNTHESIS BIFUNCTIONAL
CC PROTEIN ARGJ BETA CHAIN (BY SIMILARITY).
CC CLEAVAGE (NONHYDROLYTIC) (BY SIMILARITY).
CC SITE 178 179
CC SEQUENCE 385 AA; 4145 MW; 8A525296D0AD7C2F CRC64;

Query Match 8.4%; Score 83; DB 1; Length 385;
Best Local Similarity 21.3%; Pred. No. 4.1;
Matches 29; Conservative 23; Mismatches 44; Indels 40; Gaps 4;

OY 7 SNLSNLYTTDAS--LNGVDKLLSAVEKMLVOKAPN-----EG 45
DB 262 TKLELTISGAKSAQAARKIGKSLNSPLVTAIYGGDPMGRLLIMAVGVPEPIPEEG 321
OY 46 IEVVEGLLVLAARTSPKQVADSVIFNSFGERNVVTEGDLKKVLCAPLTFRT 105
DB 322 LQIFG-----TLPVKEANPEFLKKLSEVLKONTLSLVNVLNVTGTLMKTFGC----- 370
OY 106 NKLRTFGRTFEAYVD 121
DB 371 -----DFTKYEIE 378

RESULT 9
PCK AGRTS STANDARD; PRT; 536 AA.
ID AC Q8U94;
AC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoenolpyruvate carboxykinase (ATP) (EC 4.1.1.49) (PEP
DE carboxykinase) (phosphoenolpyruvate carboxylase) (PEPCK).
GN PCRA OR ATU0035 OR ACR_C35.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Secubal U.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendinning J., Deacher G., Gillet M., Grant C.,
RA Kutayian T., Levy R., Li M.-J., McLelland E., Palmeri A.,
RA Raymond C., Rouse G., Sasmithachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Petry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RA "The genome of the natural genetic engineer Agrobacterium tumefaciens
RA C58.";
RA Science 294:2317-2323(2001).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RX Gochner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RX Quodillo B., Goldman B.S., Cao Y., Askerazi M., Halling C., Mullin L.,
RX Hounhel K., Gordon J., Vaudin M., Iartchouk O., Bep A., Liu F.,
RX Wollam C., Allinger M., Vaudin D., Scott C., Lappas C., Marxz B.,
RX Flanagan C., Crowell C., Garsen J., Lomo C., Sear C., Strub G.,
RX Cleo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
CC -1- CATALYTIC ACTIVITY: ATP + oxaloacetate = ADP + phosphoenolpyruvate
CC + CO(2).
CC -1- PATHWAY: Rate-limiting gluconeogenic enzyme.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase (ATP)
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE008978; AAL4106.1; ALT_INIT.
CC EMBL: AE007946; AAK85859.1; ALT_INIT.
CC HAMAP: MF_00453; -; 1.
CC InterPro: IPR001272; PEPCK ATP.
CC Pfam: PF01293; PEPCK ATP; 1.
CC ProDom: PD004723; PEPCK ATP; 1.
CC TIGRFams: TIGR00224; PCRA; 1.
CC PROSITE: PS00532; PEPCK ATP; 1.
CC Gluconeogenesis; Lyase; Decarboxylase; ATP-binding; Complete proteome.
CC FM BIND 236 243
CC NP BIND 236 243
CC SEQUENCE 536 AA; 57882 MW; 6F90B8F48C86A5A CRC64;

Query Match 8.4%; Score 83; DB 1; Length 536;
Best Local Similarity 26.9%; Pred. No. 6.1;
Matches 47; Conservative 18; Mismatches 72; Indels 38; Gaps 9;

OY 20 LNGVDKLLSAVEKMLVOKAPNNGEIEVFGLL--LYLAARTS---PKVQADSVI 74
DB 237 LSGTGKTLISADPARLITDDEHGWGEGHIFVEGCGYAKAIKLSSEAEPEIYAA----- 291
OY 75 FNSNFGEE--RNVYVVEGDLKKVLDGCAPLRTFNKLRFTGRTFEAYVDCAVX-HKLP 131
DB 292 -TNRFTVLNVLNVDLSRPFDFNDN--SLTENTRS-----AYLHRTF 331
OY 132 QLNAAELGIPAEADSYLAADFLGTCPKXSEL--QOSRKFASMAVK--TEGGV 181
DB 332 NASETGAHPKTIWLTADAFGLPPIARLPPEQMYHFLSGYAKVAGTERGV 386

RESULT 10
ACKA THETN STANDARD; PRT; 401 AA.
ID ACRA THETN
AC 08R9VA;
AC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acetate kinase (EC 2.7.2.1) (Acetate kinase).
GN ACRA OR TTE1481.
OS Thermomanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermomanaerobacteriales;
OC Thermomanaerobacteriaceae; Thermomanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=MB4 / JCM 11007;
RA MEDLINE=21992816; PubMed=11997336;

```

BAO Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
 Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
 Tan H., Chen R., Wang J., Yu J., Yang H.,  
 "A complete sequence of *T. tengcongensis* genome."  
 RL Genome Res. 12:689-700 (2002).  
 CC -1- CATALYTIC ACTIVITY: ATP + acetate = ADP + acetyl phosphate.  
 CC -1- PATHWAY: Conversion of acetate to acetyl-CoA, first step.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the acetate kinase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE013105; AAM24700.1; -.  
 DR HAMAP; MF\_00020; -; 1.  
 DR InterPro; IPR000890; Acetate\_kin.  
 DR InterPro; IPR004372; AckA.  
 DR Pfam; PF00871; Acetate\_kinase; 1.  
 DR PRINTS; PR00471; ACETATEKINASE.  
 DR TIGRFAMs; TIGR00016; ackA; 1.  
 DR PROSITE; PS01075; ACETATE\_KINASE\_1; 1.  
 DR PROSITE; PS01075; ACETATE\_KINASE\_2; 1.  
 DR Transferrase; Kinase; Complete proteome.  
 KM SEQUENCE 401 AA; 44361 MW; PD71C642F1E01950 CRC64;  
 SQ  
 Query Match 8.2%; Score 81.5; DB 1; Length 401;  
 Best Local Similarity 19.3%; Pred. No. 5.9;  
 Matches 40; Conservative 35; Mismatches 87; Indels 45; Gaps 7;  
 QY 22 GVDKLLSAVY---EKMLVOKGAPN--EGIEVVFGLL-YALAAARTSPKVCADSDVIF 75  
 DB 35 GINSLSLTHQBGKQKXKIQEDMKHKEALQVLEVDGEIVGIDMEIDAVGRVHA 94  
 QY 76 SNSFGERNVVTGDLKKVLDDG---APLTRFTN-----KLRFGFTTTE 117  
 DB 95 GGEFTDSVLIDDEVIRK-LEDCIDLAPLNHPANIEGKACQIMPSPVAVFDIAFHQ 153  
 QY 118 AYVDFCLAY-----KHKLPQNAAEIGIPEDSYLAADPLGTCP 157  
 DB 154 TMPYAVYIPYPIYVEYKRRIRYRGFGTSHKRYSMAPAILGRPIELKIVTCHLNGA 213  
 QY 158 KLSLQOSRKMFASMYALKTEGVVNT 184  
 DB 214 SITAVKXGKSIDISMGFTPLEGLAMGT 240  
 RESULT 11  
 LON2 BORBU STANDARD; PRT; 813 AA.  
 ID LON2 BORBU  
 AC OS1558;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE ATP-dependent protease La homolog (EC 3.4.21.-).  
 GN BB0613.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
 NCBI\_Taxid=139;  
 [1]  
 RN SEQUENCE FROM N.A. / B31;  
 RC STRAIN=ATCC 35210 / B31;  
 RX MEDLINE=98065943; PubMed=9403685;  
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
 Rasthigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
 Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,  
 Peterson J., Kierulff A.R., Quackenbush J., Salzberg S., Hanson M.,  
 van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,  
 Uitterback T., Wetchey L., McDonald L., Artiach P., Bowman C.,

Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 Smith H.O., Venter J.C.;  
 "Genomic sequence of a Lyme disease spirochete, *Borrelia*  
*burgdorferi*."  
 RL Nature 390:580-586 (1997).  
 CC -1- SIMILARITY: Belongs to peptidase family S16.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE001162; AAC69662.1; -.  
 DR PIR; D70176; D70176.  
 DR MEROPS; S16.DPW; -.  
 DR TIGR; BB0613; -.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR003959; AAA\_ATPase\_cent.  
 DR InterPro; IPR008269; Pept\_S16\_C.  
 DR InterPro; IPR004815; Pept\_S16\_Ion.  
 DR InterPro; IPR003111; Pept\_S16\_N.  
 DR InterPro; IPR008268; Peptid\_S16\_AS.  
 DR InterPro; IPR001984; Peptidase\_S16.  
 DR Pfam; PF000004; AAA; 1.  
 DR Pfam; PF02190; LON; 1.  
 DR Pfam; PF05362; Lon\_C; 1.  
 DR PRINTS; PR00830; ENDOLAPTASE.  
 DR SMART; SM00382; AAA; 1.  
 DR SMART; SM00464; LON; 1.  
 DR TIGRFAMs; TIGR00763; lon; 1.  
 DR PROSITE; PS01046; LON\_SER; 1.  
 KM Hydrolase; Serine protease; ATP-binding; Complete proteome.  
 FT NP BIND 369 376  
 FT ACT SITE 719 719 BY SIMILARITY.  
 SQ SEQUENCE 813 AA; 92312 MW; 5129AA1498C5D0F6 CRC64;  
 Query Match 8.2%; Score 81.5; DB 1; Length 813;  
 Best Local Similarity 31.9%; Pred. No. 14;  
 Matches 37; Conservative 19; Mismatches 43; Indels 17; Gaps 6;  
 QY 8 NLSN-LVITDASLNGVCKLISAVEKMLVOKGAPNGIEVFGLLYALAAATTSKV 66  
 DB 475 DLSNLFPTTANSLNGMKRPLID-RMEIIRKYGSIYIKLAKIPLISIXSFIDKV 533  
 QY 67 Q-RADSDVIFSNFSGERNVVTG--DLKYLDDGCAPLTRFTNKLRTFGRTTAY 119  
 DB 534 YRIEDDVIFNLI---RNVTWESGVGLKRVL-----TNLRRLVRELLEY 577  
 RESULT 12  
 SP5Y MOUSE STANDARD; PRT; 366 AA.  
 ID SP5Y MOUSE  
 AC P97355;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Spermatine synthase (EC 2.5.1.22) (Spermatidne aminopropyltransferase)  
 GN SPM5Y.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_Taxid=10090;  
 [1]  
 RN SEQUENCE FROM N.A. / B31;  
 RP Strom T.M., Francis F., Lorenz B., Boedich A., Eeone M.-J.,  
 Lehersch H., Mettinger T.,  
 "Pex gene deletions in Gy and Hyp mice provide mouse models for  
 X-linked hypophosphatemia."  
 RT Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.

RN [2] SEQUENCE FROM N.A.  
 RC TISSUE=Diaphragm.  
 RA Nitraren K., Korhonen V., Janne J.,  
 RT "Nucleotide sequence of mouse spermidine aminopropyltransferase  
 RT CDNA."  
 RN Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Olfactory epithelium;  
 RX MEDLINE=2238257; PubMed=12477932.  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Altschuler R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Alekshun S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Manisera K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Tashyuk S., Carrinci F., Pirange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Paley J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shcherbko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences."  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16839-16903 (2002).  
 [4]  
 RN SEQUENCE OF 316-366 FROM N.A.  
 RP MEDLINE=9813337; PubMed=9467015;  
 RX Lorenz B., Francis F., Gempel K., Boeddrich A., Josten M., Schmahl W.,  
 RA Schmidt J., Lebrach H., Meltinger T., Strom T.M.;  
 RT "Spermine deficiency in Gy mice caused by deletion of the spermine  
 RT synthase gene."  
 RL Hum. Mol. Genet. 7:541-547 (1998).  
 CC -1- CATALYTIC ACTIVITY: S-adenosylmethioninamine + spermidine = 5'-  
 CC methylthioadenosine + spermine.  
 CC -1- PATHWAY: Biosynthesis of spermine from spermidine.  
 CC -1- SIMILARITY: Belongs to the spermidine/spermine synthase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL, Y09419; CAA70573.1; -  
 DR EMBL, AF031486; AAB86631.1; -  
 DR EMBL, BC046623; AAB86623.1; -  
 DR EMBL, AJ000093; CAA03919.1; -  
 DR EMBL, AJ000087; CAA03918.1; -  
 DR EMBL, AJ000088; CAA03918.1; JOINED.  
 DR EMBL, AJ000089; CAA03918.1; JOINED.  
 DR EMBL, AJ000090; CAA03918.1; JOINED.  
 DR EMBL, AJ000091; CAA03918.1; JOINED.  
 DR EMBL, AJ000092; CAA03918.1; JOINED.  
 DR MGD; MG1:109430; Sns.  
 DR GO; GO:0008215; P:spermine metabolism; IMP.  
 DR InterPro; IPR01045; Spermine synthase.  
 DR Pfam; PF01564; Spermine synth. 1.  
 DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.  
 KW Transferase. 132  
 FT DOMAIN 132 232  
 FT BINDING TO DECARBOXYLATED SAM  
 FT (POTENTIAL).  
 FT  
 SQ SEQUENCE 366 AA; 41313 MW; D549F319F51C43C5 CRC64;  
 Query Match 8.1%; Score 80; DB 1; Length 366;

Best Local Similarity 22.9%; Pred. No. 7.2;  
 Matches 52; Conservative 34; Mismatches 65; Indels 76; Gaps 13;  
 QY 24 DKLLSAVEKMLVQKAPNGIEVY-----FGLLVLAARTPKVGRASDVIFNSF 79  
 DB 128 DERLVEYIDVEVDSDPYQNIKILNSKQFNTL-----ILGDVLAESDLATYAI 181  
 QY 80 -----GERNVVTEGD-----LK-----KVLGCAPLTRFTNKLT 110  
 DB 182 MSGKEDYTGKQVLLGGDGGILLCEIVKLPKVTWVEIDQWYDCKKVR-----RT 236  
 QY 111 FERTTT-----EAYVDFCI-----AYHKLPQNLAAELGIPADSVLADF 152  
 DB 237 CDVDLNTKGDYQYLIDCIPVLKMKVAKEGREFDYINDLVAVPISTSPEDS--TWDF 294  
 QY 153 LGTCPTLS--ELQGRKMFASWYALKTEGVVN--TPVSNL--RQLGR 194  
 DB 295 LALIDLSMKVLKQDKYF-----TQNCVNLTEALSLVEBQLGR 334  
 RESULT 13  
 ID OTC SYNY3 STANDARD; PRT; 308 AA.  
 AC 055497;  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ornithine carbamoyltransferase (EC 2.1.3.3) (OTCase).  
 GN ARGF OR SL10902.  
 OS Synchocystis sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.  
 CX NCBI\_TaxID=1148;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96127529; PubMed=8590279;  
 RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,  
 RA Sugita M., Tabata S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synchocystis sp. strain PCC6803. I. Sequence features in the 1 Mb  
 RT region from map positions 64% to 92% of the genome."  
 RL DNA Res. 2:153-166 (1995).  
 CC -1- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate  
 CC + L-citrulline.  
 CC -1- PATHWAY: Arginine biosynthesis; sixth step.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -1- SIMILARITY: Belongs to the ATCase/OTCase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D6406; BAA10847.1; -  
 DR PIR; S76000; S76000.  
 DR HSSP; P04391; IAKM.  
 DR HAMAP; MF 01109; -; 1.  
 DR InterPro; IPR006130; Asp/Om Cotransf.  
 DR InterPro; IPR002292; Orn carbtransf.  
 DR InterPro; IPR006131; OTCase\_O.  
 DR InterPro; IPR006132; OTCase\_P.  
 DR Pfam; PF00185; OTCase; 1.  
 DR Pfam; PF02729; OTCase; 1.  
 DR PRINTS; PR00100; AOTCASE.  
 DR TIGRFAMs; TIGR00658; orn carb tr; 1.  
 DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.  
 KW Arginine biosynthesis; Transferase; Complete proteome.  
 FT SITE 33 33  
 FT IMPORTANT FOR STRUCTURAL INTEGRITY (BY  
 FT SIMILARITY).  
 FT SITE 53 57  
 FT CARBAMOYLPHOSPHATE BINDING (BY  
 FT SIMILARITY).

Query Match: 8.0%; Score 79; DB 1; Length 308;  
 Best Local Similarity 24.8%; Pred. No. 7.2; Indels 30; Gaps 7;  
 Matches 39; Conservative 25; Mismatches 63; Indels 30; Gaps 7;

FT SITE 104 104 CARBAMOYLPHOSPHATE BINDING (BY  
 FT SITE 131 131 CARBAMOYLPHOSPHATE BINDING (BY  
 FT SITE 144 144 SIMILARITY).  
 FT SITE 144 144 IMPORTANT FOR STRUCTURAL INTEGRITY (BY  
 FT SITE 264 267 ORNITHINE BINDING (BY SIMILARITY).  
 SQ SEQUENCE 308 AA; 33616 MW; C6CD2E49892CFA CRC64;

Query Match: 8.0%; Score 79; DB 1; Length 308;  
 Best Local Similarity 24.8%; Pred. No. 7.2; Indels 30; Gaps 7;  
 Matches 39; Conservative 25; Mismatches 63; Indels 30; Gaps 7;

QY 28 LSAVEKMLVQKAP-NEG1-----EVFGLLLYLAARTSPKVGADSDVIFSNFGE 81  
 DB 17 LTIEMKSLQLADLAKSGVLKPKCRKILGLLFYKASTRT--RVSFYAMTQLGQGVLD 73  
 QY 82 RNVVATE-GDLKAVLDGCAPIRTFNKL--RTGRTFTAVYDFCIAYKHKLPLQNAAE 138  
 DB 74 LNSVTQVGEGEPIQDTARVLDKYIDILAVRTFKQDTLQTFADHA---KMPITNALSD 128  
 QY 139 LGIPADSYLAADFLGTCEPLSELQOSRKMFSAMVAL 175  
 DB 129 LEHP-----CQLADLQITKECFGLBGL 152

RESULT 14  
 ID CPXC AGRTU STANDARD; PRT; 422 AA.  
 AC P2466;  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome P450-plnf1, plant-inducible (EC 1.14.-.-).  
 GN CYP103 OR PINF1 OR VIRH1.  
 OS Agrobacterium tumefaciens.  
 OG plasmid pT1A6.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
 NCBI\_TaxId=358;  
 RN NCBI\_TaxId=358;  
 RX MEDLINE=89213933; PubMed=2708311;  
 RA Kameneto R.H., Powell A.T., Akiyoshi D.E., Regier D.A.,  
 RA Kesterter R.A., Nester E.W., Hawes M.C., Gordon M.P.;  
 RT "Nucleotide sequence and analysis of the plant-inducible locus pInF  
 from Agrobacterium tumefaciens.";  
 RL J. Bacteriol. 171:2506-2512 (1989).  
 CC -1- FUNCTION: NOT ESSENTIAL FOR VIRULENCE, BUT MAY BE INVOLVED IN THE  
 CC DETOXIFICATION OF PLANT PROTECTIVE AGENTS AT THE SITE OF WOUNDING.  
 CC -1- INDUCTION: TRANSCRIPTONALLY ACTIVATED IN THE PRESENCE OF WOUNDED  
 CC PLANT TISSUE AND BY PLANT PHENOLIC COMPOUNDS, SUCH AS  
 CC ACETOSYRINONE.  
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: M19352; AAA82502.1; -.  
 DR PIR: A32306; A32306.  
 DR InterPro: IPR001128; Cytochrome\_P450.  
 DR Pfam: PF00067; P450.1.  
 DR PRINTS: PR00385; P450.  
 DR PROSITE: PS00086; CYTOCHROME\_P450; 1.  
 DR Oxidoreductase; Monooxygenase; Electron transport; Heme; Plasmid.  
 KW METAL 369 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 FT SEQUENCE 422 AA; 47519 MW; 6A9FE4AA9B7E2302 CRC64;

Query Match: 8.0%; Score 79; DB 1; Length 422;  
 Best Local Similarity 23.6%; Pred. No. 11;  
 Matches 51; Conservative 29; Mismatches 60; Indels 76; Gaps 14;

QY 20 LNVGDKLLSVEKMLVQKAPNEGIE--VFGLLLYLAARTSPKVGADSDVIFSN 77  
 DB 70 LIGDPTPTQIETELM-----NRGVKAGAVDFI-----DHSMLFSN 107  
 QY 78 --SFGERNVVVTEGDLKKVLDGCAPI-RTFT-----NKLRTGRTFTAVYDFCIAYK 127  
 DB 108 GETHGKRSGLSKAFSPRWTEALRPEIAKITECLMDLQKVDF--NFTMYAS----- 159  
 QY 128 HKLPQLAAAEIGPAD-----SYLADFLGTGCKSELC----- 163  
 DB 160 -QLPALTIASVGLPSDTPFTFLVYKVRCLSPSRDEFEIEASALELDQYRSVI 218  
 QY 164 --QSRKM--FASVYALKT--EGGVNTPVSNLRQL 192  
 DB 219 ADSGRMRDPLSRV-LKAVREAGTL-SPIEIMQL 252

RESULT 15  
 ID PPCK FUSNN STANDARD; PRT; 527 AA.  
 AC O8RE12;  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Phosphoenolpyruvate carboxylase (ATP) (EC 4.1.1.49) (PEP  
 DE carboxylase) (Phosphoenolpyruvate carboxylase) (PEPCK).  
 GN PCKA OR FN1120.  
 OS Fuschacterium nucleatum (subsp. nucleatum).  
 OC Bacteria; Fuschacteria; Fuschacteriales; Fuschacteriaceae;  
 OC Fuschacterium.  
 NCBI\_TaxId=76856;  
 RN NCBI\_TaxId=76856;  
 RX MEDLINE=2186394; PubMed=11889109;  
 RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,  
 RA Bhattacharya A., Barmen A., Gardner W., Grechkin G., Zhu L.,  
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Goldsman E., Bernal A.,  
 RA Larsen N., D'Souza W., Malinas T., Pusch G., Haselkorn R.,  
 RA Fontein M., Kyrides N., Overbeek R.;  
 RT "Genome sequence and analysis of the oral bacterium Fuschacterium  
 RT nucleatum strain ATCC 25586.";  
 RL J. Bacteriol. 184:2005-2018 (2002).  
 CC -1- CATALYTIC ACTIVITY: ATP + oxaloacetate = ADP + phosphoenolpyruvate  
 CC + CO(2).  
 CC -1- PATHWAY: Rate-limiting gluconeogenic enzyme.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the phosphoenolpyruvate carboxylase (ATP)  
 CC family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AE010616; AAL95316.1; -.  
 DR HAMAP: MF\_00453; 1.  
 DR InterPro: IPR001272; PEPCK\_ATP.  
 DR Pfam: PF01293; PEPCK\_ATP; 1.  
 DR ProDom: PD004723; PEPCK\_ATP; 1.  
 DR TIGRFAMs: TIGR00224; PCKA; 1.  
 DR PROSITE: PS00532; PEPCK\_ATP; 1.  
 DR Gluconogenesis; Lyase; Decarboxylase; ATP-binding; Complete proteome.  
 KW NP\_BIND 237 ATP (BY SIMILARITY).  
 FT SEQUENCE 527 AA; 59055 MW; 275849FD254AC01 CRC64;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 5, 2004, 17:50:07 ; Search time 45 Seconds  
(Without alignments)  
1388.280 Million cell updates/sec

Title: US-09-613-486-15  
Sequence: 1 MEI MSDSNLSNLVITDASL.....GVVNTPVSNLRQLGRREVM 198

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_plant:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriophage:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	991 100.0	198 12	071213	071213 grapevine 1
2	984 99.3	198 12	039857	039857 grapevine 1
3	921 92.3	198 12	08BE23	08BE23 grapevine 1
4	297 30.0	204 12	09Q629	09Q629 sugar beet
5	296 29.9	204 12	09Q664	09Q664 sugar beet
6	295 29.8	204 12	089910	089910 sugar beet
7	293 29.6	204 12	003724	003724 sugar beet
8	286 28.9	204 12	008531	008531 sugar beet
9	265 26.9	215 12	065858	065858 beet yellow
10	130.5	223 12	08JY1	08JY1 citrus tris
11	129.5	223 12	09Q8C4	09Q8C4 citrus tris
12	125.5	222 12	0990F0	0990F0 citrus tris
13	125.5	223 12	0990B2	0990B2 citrus tris
14	125.5	223 12	09Q8G5	09Q8G5 citrus tris
15	125	223 12	08B373	08B373 citrus tris
16	124.5	223 12	09Q8G2	09Q8G2 citrus tris

17	123.5	223 12	09E7M2	09E7M2 citrus tris
18	123.5	223 12	09E7M4	09E7M4 citrus tris
19	123.5	223 12	09PXR0	09PXR0 citrus tris
20	123.5	223 12	09Q8G3	09Q8G3 citrus tris
21	123	212 12	09DMX1	09DMX1 citrus tris
22	123	223 12	09E7M3	09E7M3 citrus tris
23	120.5	195 12	08V1P7	08V1P7 citrus tris
24	120.5	222 12	0990N9	0990N9 citrus tris
25	120.5	223 12	09Q8G1	09Q8G1 citrus tris
26	120.5	223 12	089948	089948 citrus tris
27	120.5	223 12	0990P1	0990P1 citrus tris
28	120.5	223 12	09M1D9	09M1D9 citrus tris
29	120.5	223 12	091EW4	091EW4 citrus tris
30	120.5	223 12	006192	006192 citrus tris
31	120.5	223 12	09PXR3	09PXR3 citrus tris
32	118.5	223 12	08JYU9	08JYU9 citrus tris
33	117.5	223 12	09PXR9	09PXR9 citrus tris
34	117.5	223 12	0991G6	0991G6 citrus tris
35	116	223 12	09Q8G6	09Q8G6 citrus tris
36	115.5	195 12	08V1P5	08V1P5 citrus tris
37	115.5	223 12	08JYV0	08JYV0 citrus tris
38	115	223 12	09PXR1	09PXR1 citrus tris
39	114.5	223 12	09DTR8	09DTR8 citrus tris
40	114.5	223 12	010474	010474 citrus tris
41	114.5	223 12	08B3R9	08B3R9 citrus tris
42	113.5	194 12	08V1P6	08V1P6 citrus tris
43	111.5	223 12	08JYU8	08JYU8 citrus tris
44	110.5	223 12	09PXR2	09PXR2 citrus tris
45	109	200 12	099AT6	099AT6 citrus tris

## ALIGNMENTS

## RESULT 1

ID 071213 PRELIMINARY; PRT; 198 AA.

AC 071213;  
DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)  
DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)  
DE 22 kDa coat protein.  
OS Grapevine leafroll-associated virus 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;  
OC Closterovirus.  
OX NCBI\_TaxId=64003;  
RN [1]  
RX MEDLINE=98264507; PubMed=9603345;  
RA Zhu H.Y., Ling K.S., Gosczyński D.E., McFerson J.R., Gonsalves D.;  
RT "Nucleotide Sequence and Genome Organization of Grapevine Leafroll-  
RT Associated Virus-2 are Similar to Beet Yellow Virus, the  
RT Closterovirus Type Member";  
RL J. Gen. Virol. 79:1289-1298(1998)  
DR EMBL: AF039204; AAC40861.1; -;  
DR GO: GO:0019028; C: viral capsid; IEA.  
DR GO: GO:0005198; F: structural molecule activity; IEA.  
DR InterPro: IPR002679; C: closter coat.  
DR Pfam: PF01785; C: closter\_coat; 1.  
KW Coat protein.  
SQ SEQUENCE 198 AA; 21660 MW; 8ED85EF2E1DFB03 CRC64;

Query Match 100.0%; Score 991; DB 12; Length 198;  
Best Local Similarity 100.0%; Pred. No. 1.3e-88;  
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MEI MSDSNLSNLVITDASLNGVDKLLSAEYKMLVCGAGNNEGIEVFGLLVALAAR	60
DB	1	MEI MSDSNLSNLVITDASLNGVDKLLSAEYKMLVCGAGNNEGIEVFGLLVALAAR	60
QY	61	TTSPKTVQADSDVIFSNSTFGERNVVVTBGLDKKVIDGCAPLIRFTNKLTFRGTTEAVV	120
DB	61	TTSPKTVQADSDVIFSNSTFGERNVVVTBGLDKKVIDGCAPLIRFTNKLTFRGTTEAVV	120

QY 121 DFCIAYKHKL.POLNMAAELGIPADSSTYLAADFLGTCTKLSLQSRKMFASMYALKTEGG 180  
 DB 121 DFCIAYKHKL.POLNMAAELGIPADSSTYLAADFLGTCTKLSLQSRKMFASMYALKTEGG 180  
 QY 181 VVNTPVSNLRQLGRREVM 198  
 DB 181 VVNTPVSNLRQLGRREVM 198

## RESULT 2

039857 PRELIMINARY: PRT: 198 AA.  
 AC 039857  
 DT 01-JAN-1998 (Tremblrel. 05, Created)  
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE Coat protein.  
 OS Grapevine leafroll-associated virus 2.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;  
 OC Closterovirus.  
 OX NCBI\_TaxID=64003;  
 RN (1)  
 RN SEQUENCE FROM N.A.  
 RA Abou-Ghanem N.;  
 RT "The nucleotide sequence of the 3' terminal region of grapevine  
 RT leafroll associated closterovirus 2.";  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Y14131; CAA74366.1; -  
 DR GO; GO:0019028; C:Viral capsid; IEA.  
 DR GO; GO:0005198; F:Structural molecule activity; IEA.  
 DR InterPro; IPR002679; Closter\_coat.  
 DR Pfam; PF01785; Closter\_coat; 1.  
 SQ SEQUENCE 198 AA; 21630 MW; 15D631E408BD3F01 CRC64;

Query Match 99.3%; Score 984; DB 12; Length 198;  
 Best Local Similarity 99.5%; Pred. No. 6,3e-88;  
 Matches 197; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MELMSDSNLSNIVITDASSLNGVDKLLSAEYKMLVKGAPNBSGIEVFGILLYALAA 60  
 DB 1 MELMSDSNLSNIVITDASSLNGVDKLLSAEYKMLVKGAPNBSGIEVFGILLYALAA 60  
 QY 61 TTSPPVQRADSDVIFSNFSGERNVVTGGDKLVLDGCAPLTRFNKLRTRGRTTEAYV 120  
 DB 61 TTSPPVQRADSDVIFSNFSGERNVVTGGDKLVLDGCAPLTRFNKLRTRGRTTEAYV 120  
 QY 121 DFCIAYKHKL.POLNMAAELGIPADSSTYLAADFLGTCTKLSLQSRKMFASMYALKTEGG 180  
 DB 121 DFCIAYKHKL.POLNMAAELGIPADSSTYLAADFLGTCTKLSLQSRKMFASMYALKTEGG 180  
 QY 181 VVNTPVSNLRQLGRREVM 198  
 DB 181 VVNTPVSNLRQLGRREVM 198

## RESULT 3

088BP3 PRELIMINARY: PRT: 198 AA.  
 ID 088BP3  
 AC 088BP3;  
 DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE Coat protein.  
 OS Grapevine rootstock stem lesion associated virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;  
 OC Closterovirus.  
 OX NCBI\_TaxID=167634;  
 OX (1)  
 RN SEQUENCE FROM N.A.  
 RA Zhang Y.; Rowhani A.;  
 RT "Nucleotide sequence of grapevine rootstock stem lesion associated  
 RT virus.";

RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF314063; AAN63472.1; -  
 DR GO; GO:0019028; C:Viral capsid; IEA.  
 DR GO; GO:0005198; F:Structural molecule activity; IEA.  
 DR InterPro; IPR002679; Closter\_coat.  
 DR Pfam; PF01785; Closter\_coat; 1.  
 SQ SEQUENCE 198 AA; 21620 MW; 370E95A9369734A9 CRC64;

Query Match 92.9%; Score 921; DB 12; Length 198;  
 Best Local Similarity 90.9%; Pred. No. 9.1e-82;  
 Matches 180; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

QY 1 MELMSDSNLSNIVITDASSLNGVDKLLSAEYKMLVKGAPNBSGIEVFGILLYALAA 60  
 DB 1 MELMSDSNLSNIVITDASSLNGVDKLLSAEYKMLVKGAPNBSGIEVFGILLYALAA 60  
 QY 61 TTSPPVQRADSDVIFSNFSGERNVVTGGDKLVLDGCAPLTRFNKLRTRGRTTEAYV 120  
 DB 61 TTSPPVQRADSDVIFSNFSGERNVVTGGDKLVLDGCAPLTRFNKLRTRGRTTEAYV 120  
 QY 121 DFCIAYKHKL.POLNMAAELGIPADSSTYLAADFLGTCTKLSLQSRKMFASMYALKTEGG 180  
 DB 121 DFCIAYKHKL.POLNMAAELGIPADSSTYLAADFLGTCTKLSLQSRKMFASMYALKTEGG 180  
 QY 181 VVNTPVSNLRQLGRREVM 198  
 DB 181 VVNTPVSNLRQLGRREVM 198

## RESULT 4

090629 PRELIMINARY: PRT: 204 AA.  
 ID 090629  
 AC 090629;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
 DE Major capsid protein.  
 OS Sugar beet yellow virus (SBYV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;  
 OC Closterovirus.  
 OX NCBI\_TaxID=12161;  
 RN (1)  
 RN SEQUENCE FROM N.A.  
 RA STRAIN=BYV-4;  
 RX MEDLINE=20079557; PubMed=10611288;  
 RA Peremyshlov V.V.; Haglatare Y.; Dolja V.V.;  
 RT "SBYV homolog functions in cell-to-cell movement of a plant virus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:14771-14776 (1999).  
 RN (2)  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=BYV-4;  
 RA Peremyshlov V.V.; Haglatare Y.; Alzhanova D.; Dolja V.V.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF190581; AAF14505.1; -  
 DR GO; GO:0019028; C:Viral capsid; IEA.  
 DR GO; GO:0005198; F:Structural molecule activity; IEA.  
 DR InterPro; IPR002679; Closter\_coat.  
 DR Pfam; PF01785; Closter\_coat; 1.  
 SQ SEQUENCE 204 AA; 22419 MW; FD6337E1D2490D33 CRC64;

Query Match 30.0%; Score 297; DB 12; Length 204;  
 Best Local Similarity 37.6%; Pred. No. 8.6e-21;  
 Matches 71; Conservative 33; Mismatches 84; Indels 2; Gaps 2;

QY 7 SNLSNIVITDASSLNGVDKLLSAEYKMLVKGAPNBSGIEVFGILLYALAAKRTSPKY 66  
 DB 11 ATPEVSLADQTCIHGBDCDKLRDFEBCIKLKGVEDKIGLALGCLVSCATIGTSNKV 70  
 QY 67 QRADSDVIFSNFSGERNVVTGGDKLVLDGCAPLTRFNKLRTRGRTTEAYVFCIA 125  
 DB 71 SVQPTSTFIKASFGSGELFLTHGELRSFLDSQKLBGKPNKLCFCRTIQKOYISPAKE 130  
 QY 126 YKHKL.POLNMAAELGIPADSSTYLAADFLGTCTKLSLQSRKMFASMYALKTEGGVNTP 185



```

Db      131 YGRLPPIRANRNGLPADHDYLAADFIISTETLTDLQCGRLILARENATHTEFS-SESP 189
QY      186 VSNLRQLGR 194
Db      190 VTSLKQLGR 198

RESULT 5
Q96664    PRELIMINARY; PRT; 204 AA.
ID        Q96664
AC        Q96664;
DT        01-FEB-1997 (TRENBLrel. 02, Created)
DT        01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT        01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE        Coat Protein (Fragment).
OS        Sugar beet yellow virus (SBYV).
OC        Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
OC        Closterovirus.
OX        NCBI_TaxID=12161;
RN        [1]
RP        SEQUENCE FROM N.A.
RC        STRAIN=California BYV-CA;
RA        Creamer R., He X.H., Yang C.H., Grantham G.;
RT        "Characterization of the 3'-proximal encoded proteins of beet yellows
RT        closterovirus."
RL        Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR        EMBL; U71295; AAB17001.1; -.
DR        GO; GO:0019028; C:Viral capsid; IEA.
DR        GO; GO:0005198; F:Structural molecule activity; IEA.
DR        InterPro; IPR002679; Closter_coat.
DR        Pfam; PF01785; Closter_coat; 1.
FT        NON_TER
SQ        SEQUENCE 204 AA; 22419 MW; D082C33EC93883B CRC64;

Query Match      29.8%; Score 296; DB 12; Length 204;
Best Local Similarity 37.6%; Pred. No. 1.1e-20;
Matches 71; Conservative 32; Mismatches 84; Indels 2; Gaps 2;

QY      7 SNLSNLTVDASSLNGVDKLLSAEVEKMLVQKAPNEGLEVFGLILYALAARTSPKY 66
Db      11 ATEFNVSIVDQTCIHGEDCDKLRNPFECIKLKGVPEDKGLALGLCLVSCATIGTSNKV 70
QY      67 QRADSDVIFNSPGE-RNVVVTGDDLKVLDCGAPLRTFKRLRTFGRTTEAYVDFCIA 125
Db      71 SVQPTSTRTKASFSQSGKELFTTHGELRSFLDSQCLBGRKNKLCFRTTQKDIYISPAKE 130
QY      126 YKHKLQPLNMAAELGIPADSDYLAADFLGTCPKLSLQSGRKMPASMYALKTEGGVNTP 185
Db      131 YRGRLPPIRANRNGLPADHDYLAADFIISTETLTDLQCGRLILARENATHTEFS-SDSP 189
QY      186 VSNLRQLGR 194
Db      190 VTSLKQLGR 198

RESULT 6
Q89910    PRELIMINARY; PRT; 204 AA.
ID        Q89910
AC        Q89910;
DT        01-NOV-1998 (TRENBLrel. 08, Created)
DT        01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT        01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE        Major coat protein.
OS        Sugar beet yellow virus (SBYV).
OC        Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
OC        Closterovirus.
OX        NCBI_TaxID=12161;
RN        [1]
RP        SEQUENCE FROM N.A.
RC        STRAIN=California;
RA        MEDLINE=98285746; PubMed=9621048;
RA        Peremylov V.V., Haglwa Y., Dolja V.V.;

```

```

RT      "Genes required for replication of the 15.5-kilobase RNA genome of a
RT      plant closterovirus."
RT      J. Virol. 72:5870-5876 (1998).
RN      [2]
RP        SEQUENCE FROM N.A.
RC        STRAIN=California;
RA        Dolja V., Peremylov V., Haglwa Y.;
RL        Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR        EMBL; AF056575; AAC25120.1; -.
DR        GO; GO:0019028; C:Viral capsid; IEA.
DR        GO; GO:0005198; F:Structural molecule activity; IEA.
DR        InterPro; IPR002679; Closter_coat.
DR        Pfam; PF01785; Closter_coat; 1.
KW        Coat protein.
SQ        SEQUENCE 204 AA; 22446 MW; D913184ECC93883B CRC64;

Query Match      29.8%; Score 295; DB 12; Length 204;
Best Local Similarity 37.6%; Pred. No. 1.4e-20;
Matches 71; Conservative 31; Mismatches 85; Indels 2; Gaps 2;

QY      7 SNLSNLTVDASSLNGVDKLLSAEVEKMLVQKAPNEGLEVFGLILYALAARTSPKY 66
Db      11 ATEFNVSIVDQTCIHGEDCDKLRNPFECIKLKGVPEDKGLALGLCLVSCATIGTSNKV 70
QY      67 QRADSDVIFNSPGE-RNVVVTGDDLKVLDCGAPLRTFKRLRTFGRTTEAYVDFCIA 125
Db      71 SVQPTSTRTKASFSQSGKELFTTHGELRSFLDSQCLBGRKNKLCFRTTQKDIYISPAKE 130
QY      126 YKHKLQPLNMAAELGIPADSDYLAADFLGTCPKLSLQSGRKMPASMYALKTEGGVNTP 185
Db      131 YRGRLPPIRANRNGLPADHDYLAADFIISTETLTDLQCGRLILARENATHTEFS-SESP 189
QY      186 VSNLRQLGR 194
Db      190 VTSLKQLGR 198

RESULT 7
Q003724    PRELIMINARY; PRT; 204 AA.
ID        Q003724
AC        Q003724;
DT        01-NOV-1996 (TRENBLrel. 01, Created)
DT        01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT        01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE        Coat protein.
OS        Sugar beet yellow virus (SBYV).
OC        Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
OC        Closterovirus.
OX        NCBI_TaxID=12161;
RN        [1]
RP        SEQUENCE FROM N.A.
RC        STRAIN=Broom's barn;
RA        Brunstedt J., Moseley J., Hull R.;
RT        "Nucleotide sequence of cDNA encoding the coat protein of beet yellows
RT        virus."
RL        Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.
RN        [2]
RP        SEQUENCE FROM N.A.
RC        STRAIN=German;
RA        Agrankovsky A.A., Koonin E.Y., Boyko V.P., Mais E., Lunina N.A.;
RA        Atabekov J.G.;
RT        "Beet yellows closterovirus: complete genome structure and
RT        identification of a leader papain-like chitin protease."
RL        Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
DR        EMBL; M59452; AAB12955.1; -.
DR        EMBL; X73475; AAB1861.1; -.
DR        PIR; S27900; S27900.
DR        GO; GO:0019028; C:Viral capsid; IEA.
DR        GO; GO:0005198; F:Structural molecule activity; IEA.
DR        InterPro; IPR002679; Closter_coat.
DR        Pfam; PF01785; Closter_coat; 1.
SQ        SEQUENCE 204 AA; 22388 MW; 92B2ADCA77543D9 CRC64;

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - nucleic search, using frame\_plus.p2n model

Run on: April 7, 2004, 01:15:49 ; Search time 2524 Seconds  
(Without alignments)  
3400.127 Million cell updates/sec

Title: US-09-613-486-15  
Perfect score: 991  
Sequence: 1 MEMSDSNLSNLTVDASSL.....GGVNTPVSNLRQLGRREV 198

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US09613486/rnatc.05042004.155121.28262/app.query.fasta\_1.391  
-DB=GenBdb1 -OPMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=ppct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFM=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09613486 @CGN 1.1\_2372.@rnatc.05042004.155121.28262 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEOUTRY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : GenBdb1.\*  
1: gb\_hsa.\*  
2: gb\_hhg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_cv.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pt.\*  
10: gb\_ro.\*  
11: gb\_stb.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vt.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pac.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vt.\*  
30: em\_hhg\_hum.\*  
31: em\_hhg\_inv.\*  
32: em\_hhg\_other.\*  
33: em\_hhg\_mus.\*  
34: em\_hhg\_pln.\*  
35: em\_hhg\_rtd.\*  
36: em\_hhg\_mam.\*  
37: em\_hhg\_vtc.\*  
38: em\_sy.\*  
39: em\_hhg\_hum.\*  
40: em\_hhg\_mus.\*  
41: em\_hhg\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991	100.0	597	6 AR138306	AR138306 Sequence
2	991	100.0	15000	14 AF039204	AF039204 Grapevine
3	991	100.0	15500	6 AR138299	AR138299 Sequence
4	984	99.3	8590	14 GLA4131	V14131 Grapevine 1
5	921	92.9	15527	14 AF314061	AF314061 Grapevine
6	297	30.0	15468	14 AF190581	AF190581 Beet yell
7	296	29.9	612	14 BYU71295	U71295 Beet yellow
8	295	29.8	15468	14 AF056575	AF056575 Beet yell
9	293	29.6	2724	14 BYVCPG	MS9452 Beet yellow
10	293	29.6	5980	14 BYVMPA	X73475 Beet yellow
11	286	28.9	6746	14 CLBY3PH	X53462 Beet yellow
12	286	28.9	15480	14 BYVDA	X73476 Beet yellow
13	285	28.8	783	6 A41914	A41914 Sequence 1
14	266.5	26.9	10545	14 BYU51931	U51931 Beet yellow
15	130.5	13.2	672	14 AF501867	AF501867 Citrus tr
16	129.5	13.1	672	14 AF184115	AF184115 Citrus tr
17	125.5	12.7	672	14 AF184114	AF184114 Citrus tr
18	125.5	12.7	672	14 AF342894	AF342894 Citrus tr
19	125.5	12.7	672	14 AF342890	AF342890 Citrus tr
20	125	12.6	1080	14 CTVCAT	AF342895 Citrus tr
21	125	12.6	19293	14 AY170468	AY170468 Citrus tr
22	125	12.6	19296	14 AY340974	AY340974 Citrus tr
23	125	12.6	19296	14 CTU16304	U6304 Citrus tris
24	124.5	12.6	672	14 AF184117	AF184117 Citrus tr
25	123.5	12.5	672	14 AF184116	AF184116 Citrus tr
26	123.5	12.5	672	14 AF220502	AF220502 Citrus tr
27	123.5	12.5	672	14 AF220504	AF220504 Citrus tr
28	123	12.4	672	14 CTR297702	CT297702 Citrus tr
29	123	12.4	672	14 AF220503	AF220503 Citrus tr
30	120.5	12.2	589	14 AF456771	AF456771 Citrus tr
31	120.5	12.2	672	14 AF342895	AF342895 Citrus tr
32	120.5	12.2	672	14 AF184118	AF184118 Citrus tr
33	120.5	12.2	672	14 AF342891	AF342891 Citrus tr
34	120.5	12.2	672	14 AF342892	AF342892 Citrus tr
35	120.5	12.2	672	14 AF342893	AF342893 Citrus tr
36	120.5	12.2	672	14 CTVCAT	U12175 Citrus tris
37	120.5	12.2	19226	14 CTU56902	U56902 Citrus tris
38	120.5	12.2	19226	14 AF260651	AF260651 Citrus tr
39	120.5	12.2	19259	14 CTU18440	U18440 Citrus tris
40	118.5	12.0	672	14 AF501869	AF501869 Citrus tr
41	117.5	11.9	587	14 AF220505	AF220505 Citrus tr
42	117.5	11.9	672	14 AF339088	AF339088 Citrus tr
43	116	11.7	589	14 AF184113	AF184113 Citrus tr
44	115.5	11.7	589	14 AF456773	AF456773 Citrus tr
45	115.5	11.7	672	14 AF501868	AF501868 Citrus tr

RESULT 1

## ALIGNMENTS

AR138306  
 LOCUS AR138306 597 bp DNA linear PAT 16-JUN-2001  
 DEFINITION Sequence 14 from patent US 6197948.  
 ACCESSION AR138306  
 VERSION AR138306.1 GI:14479815  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE  
 1 (bases 1 to 597)  
 Zhu, H.-Y., Ling, K.-S. and Gonsalves, D.  
 Grapevine leafroll virus (type 2) proteins and their uses  
 JOURNAL Patent: US 6197948-A 14 06-MAR-2001;  
 FEATURES  
 source  
 1..597  
 /organism="Unknown"  
 /mol\_type="unassigned DNA"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,39e-97 Length: 597  
 Score: 991.00 Matches: 198  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-613-486-15 (1-198) x AR138306 (1-597)

QY 1 MetGluLeuMetSerAspSerAsnLeuSerValIleThrAspAlaSerSerLeu 20  
 Db 1 ATGAGTGGATGTCGACGACGACCTTACCACTGGATGATACCGACGCTTACGCTA 60  
 QY 21 AengGlyValAspGlyValLeuSerValGluValGluValMetLeuValGluValGly 40  
 Db 61 AATGTTGTCGACGACGACCTTATCTGCTGATGATGATGATGATGATGATGATGATG 120  
 QY 41 AlaProAengGluGlyIleGluValValPheGlyLeuLeuLeuValAlaValAlaArg 60  
 Db 121 GCTCTTACGAGGATGATGAGAGTGTGCTGCTTACCTTACGACCTGCGGCAAGA 180  
 QY 61 ThrThrSerProGlyValGluValAlaAspSerAspValIlePheSerAspSerPheGly 80  
 Db 181 ACCAGGTCCTCTTACGAGGTCGACGACGATTCAGCGCTTATATTTCAATAGTTTGGGA 240  
 QY 81 GluAcGanValValIleThrGluGlyAspLeuValValLeuAspGlyCysAlaPro 100  
 Db 241 GAGAGGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 300  
 QY 101 LeuThrArgPheThrAsnLeuValGluValPheGlyValGluValPheThrGluValAla 120  
 Db 301 CTCCTAGGTCCTTACCACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 360  
 QY 121 AspPheCysIleAlaTyrllyHisIlySerLeuProGluLeuAanAlaAlaGluLeuGly 140  
 Db 361 GACTTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 420  
 QY 141 IleProAlaGluAspSerTyrlleuAlaAlaAspLeuGlyThrCysProGlyLeuSer 160  
 Db 421 ATTCAGGTCGAGATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 480  
 QY 161 GluLeuGluGluSerArgIlySerPheAlaSerMetTyrlAlaLeuValThrGluGly 180  
 Db 481 GATTTACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 540  
 QY 181 ValValAspThrProValSerAsnLeuValGluValGluValGluValGluValMet 198  
 Db 541 GTGCTAAATACACACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 594

RESULT 2  
 AF039204 15000 bp RNA linear VRL-22-MAY-1998  
 LOCUS AF039204  
 DEFINITION Grapevine leafroll-associated virus 2 methyltransferase/helicase

polyprotein gene, partial cds; and RNA-dependent RNA polymerase, putative transmembrane small hydrophobic protein, 65 kDa chaperone protein, 63 kDa protein, 25 kDa diverged coat protein, 22 kDa coat protein, 19 kDa protein, and 24 kDa protein genes, complete cds.  
 AF039204.1 GI:1323909  
 ACCESSION AF039204  
 VERSION AF039204.1  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE  
 1 (bases 1 to 15000)  
 Zhu, H.-Y., Ling, K.-S., Gosczyński, D.E., McFerson, J.R. and Gonsalves, D.  
 Nucleotide sequence and genome organization of grapevine leafroll-associated virus-2 are similar to beet yellows virus, the closterovirus type member  
 J. Gen. Virol. 79 (Pt 5), 1289-1298. (1998)  
 MEDLINE  
 PUBMED 9624507  
 JOURNAL 9603345  
 TITLE  
 AUTHORS  
 REFERENCE  
 2 (bases 1 to 15000)  
 Zhu, H.-Y., Ling, K.-S. and Gonsalves, D.  
 Direct Submission  
 Submitted (18-DEC-1997) Plant Pathology, Cornell University, New York State Agricultural Experiment Station, Geneva, NY 14456, USA  
 FEATURES  
 source  
 1..15000  
 /organism="Grapevine leafroll-associated virus 2"  
 /mol\_type="genomic RNA"  
 /db\_xref="taxon:64003"  
 <..7423  
 /function="replication and unknown functions"  
 /note="larger than 277 kDa; ORF1a; contains domains for two papain-like leader proteases, a methyltransferase and a helicase; identified by sequence comparison"  
 /codon\_start=2  
 /product="methyltransferase/helicase polyprotein"  
 /protein\_id="AAC40855.1"  
 /db\_xref="GI:1323910"  
 /translation="ADVYAMLRVYCGKFPVLNRSVITPDGRCYLAHMYLCAFYCR  
 /PRESIDYALGMPYVLRACVKNFVAGAGILARGYTSRNVHVCYDAAYVYFR  
 NLSGRIGGSPDSTLSTVITVYKLSGLGGLPKNTAFACDILRVNLPVSGGQSS  
 VKTKEDHRTVEERAGSVYQPOKSIDERGCGRVSGFSLVNLNLEVRKVA  
 AGLLRVYQGDMDPFRSFTQAGHLLVSSSSRVCLELSPSNFLRYVLPSCGD  
 YAMPSFPAAGRFPLVMTIRITVPGFCYLAHRCACAPKILRGPRPKFDIYGAFTAA  
 KLRNWSSELGSRSLVNLVGYATSRVPHCDYDKPTKDLRLMSAVTAGDQGEVY  
 PSDITPAMKQKTIKAVYDLVYGTSLSLKSLIEKQIDFKDVOSLKORFIVKVPY  
 MEATQNSLTRFYPQPELKFSSHSRSDHPAASRLBNELVRLCGNSVSDIGCPL  
 FLHAKTORVAVCRPVLVDGKQARVVDQYSNVRLGMDTLEGRNVDICHPPL  
 GACHSSAMMVQVYDASLYEIGAMIKKSRITVLMVTPGSELDREREVMSLD  
 CTEIVDVAADVVMYKFGSSCYSHKSLIKMTITPTLVLGGFLFVEMYEVAMGVY  
 KITKSEVSSICTLTLKRRRNSDVVYVYKIPRPKRMLPGDITVYLSKFSRV  
 FYVVCNCSAVSKTFEYWSFTKSKSRVITSGKTIHKDYNLDKTYESPAAVNLAS  
 GVSRLASVYAKNLSHSGDSPTLEASVFLREKIRNMTLNFNRLQVLRVAFAT  
 LVSFLDLDTLESITDFAECVVALELDELGLRAENKRLRNLAGSIAKLSAI  
 VVDIDSKPSPKQVGNSSSENADREVQPGIARGSNVGVGFLFHVDSALRLKRYA  
 TPOORIKSVYRFLDSAVFLDYNQNLFFILFLESGYSCMPAFIANKGDSLRVSA  
 VCAVREVAATSCANASVSKAKMTITFAAVCAMFNSCGSPGSGREYKSIHYTVGLF  
 DITPEPDSXYLPIEYLSNAICGATVYLSGSSILNMFILQITGFSLEVVAVRVR  
 VTHGLSTITDVIKGVFQIYSHLNLNMGVAAQSAFSLVGLVPLVKKCVSLIFIL  
 REDITSGFTRHGISFPLSLTLKFLKGLVDELSSITGGVDSKHKVKKATQAR  
 TTVMOYPAVAVDALSKAKIYNNTSRRTFKEDGSSSDGACSEYSCDBSGGLK  
 GGSSYGFSLIAFFSRIMMGARLIVKVKECGKLFEFLSLHSEFRVFGKNTDV  
 GYVDFLPTGIVETLSIECDQIELLGDLDGDDASITDMAYEFSEDFLASIEEP  
 PFAKRGKSNLALIALIEYANLPRIVYSSKSPDLFAFAKELSLIEKKEVFR  
 KGLVAVRYETORFLRSRMLGANNVVSFPDLPALMKRVSGLFASVYRPL  
 RGSYWCVAERERKFRALVCLIGSLPFRVIVAKKCEELVSSARRYEIKIFLR  
 QKTVLSLNFCHLPSDDVDSASAGLKGASRMTLPHLYVLAASALISGEGKTL  
 LSHHNLFLCFALVDVNVLLIVLGGLSFFVPITSLFAMLLQDPRFVEYBEKLVTA  
 PEFFLKSPRAPLKGFPECVANSTVSTVRLRLCFVKMLKLRGRLVAGLGGH  
 ROKAVAVIPSNRVVTDGVERLSVKMGVVALTEELIEDDLSAVIEKLRNRNDTN  
 DDEFTRPAHEQMOEVTTFCSKANSAGLALERVALVEDADIKSEKLSKTVENVRKOSTT  
 SEEVAVALDDDEAVEEISVADBRDSEFKTVISEYLNRLNSSFEPKPIVVDNDGTG

CDS  
 GLTNVAREFYMOELALFEIHSKLCITYOQLRVNFDNSVAPCSEDAOLYKRNKSTI  
 VQKEVRHLIKQHDHDFEDGKISINKRRGGVLYHDNLAFVLSNLSLAPYPSRS  
 FVETNSVILLYEAPPGGKTYTLIDISFLKFKGSEVTMLITANKSSQVEILKRE  
 KEVNIIECQKRDKRSPPKSIYITIDAYLMHNRGCDADVLFIDECFWHAGSVLACIEF  
 TRCHVIMFGDSDROIYIERNELDKCYGLDFVLOCRVGNISYCPMDYCAMLS  
 TVYGNJIAITYKGBSGKSMRINEISNDVDVADVGSTFCMLQSEKLEISGPIKRG  
 LITLANTVTHEAGETTYARINIVRLKRODEPKSRHITVVALSPRTSDITVYLAAR  
 RGDATCAIQKAEIVNKRKRVPTSGSGITININVKVDENDSRCAASAPISVINDF  
 LNEVPGTAVIDGDLSDPSTGPECGASGIVVRDNISSTITDHDQRV"  
 <7422..8801  
 /function="replication"  
 /note="RBDP; 52 kDa; similar to RNA polymerases of other  
 closteroviruses; presumably expressed via +1 ribosomal  
 frameshift."  
 /codon\_start=1  
 /product="RNA-dependent RNA polymerase"  
 /protein\_id="AAC40856.1"  
 /db\_xref="GI:3123911"  
 /translation="SVYRQAIPIRRKPSLOENLYSFEARNVNSTCDRNTASMEGEA  
 MAMNCLRCRCDLAPSLRDVLSITRSGIEOMLEKRTPSQIKALMKDYESPLEIDE  
 ICRFLKLVKRDAYKLDSSCLTGHSAONIMFERKSNALIESPIFNEVNRIMCCIKP  
 NIKFTFEMTRDFASVSNMGLDDVYHICEVDPSKIDKSDAFVAFEBEYWKELGV  
 DEBLALIMCGERLSIANTLDQSLFTIENQKSGASNWIGNSLVTLLISLYDVR  
 NFPAIYISGDSLIIFSRSEISNYADICTDMGEFTKEMSPVPCSKRVLVWCGHFT  
 FVPDPYKLFVKLGAVKEDVMDPLFETFTSFKDITSDFNDERLIQRLAIVALKVEVQ  
 TGNUTLALSVIHCIRSNFLSFKLYPRVKMGQVFTSVKALKLSGCSLFSFMTFPG  
 QAWMWDE"

CDS  
 8865..9035  
 /note="6 kDa; probably membrane-associated; similar to  
 small hydrophobic proteins of other closteroviruses"  
 /codon\_start=1  
 /product="putative transmembrane small hydrophobic  
 protein"  
 /protein\_id="AAC40857.1"  
 /db\_xref="GI:3123912"  
 /translation="MNVQLQPECLFLNLAVFATPIFILVPRVILKSRQKHENAV  
 PVVGGGFSTVY"  
 9051..10850  
 /note="p65; HSP70; similar to heat shock 70 proteins;  
 identified by sequence comparison"  
 /codon\_start=1  
 /product="65 kDa chaperone protein"  
 /protein\_id="AAC40858.1"  
 /db\_xref="GI:3123913"  
 /translation="MNVFGLDPGTTPSTVCVYKGRVFSKONNSAYIPTLYLFSDS  
 LNHFTVGEASIMSNLKYKGSFYADLKRWQCDSDADLIDLKHYSVRLKIGSG  
 LNETVSYISGFGTYSKSAHLPGILALFIKAVISCAEGAFCTGTCISVPAYDSQ  
 NPTDCCVSLSGVQCVNINEPSAALISAGNSIGKSNALAVDFGGTDPVILISYR  
 NMTFVPAASGDNLGSRDVPRAFLTLFSLTIDPILTDINIKESLTKDPAETVY  
 TLRGVDRKEDVRVKNKILTSVMLPYNRILTKILESTKYSASANSASRAVKKDYLVI  
 GGSSYLPLGLADVLTKQSVDRILKSDPRAVAVGCLVSCSGSGGLLILCAHAT  
 VALIDRCHQIICAPAGAPLIPFSGMPLYLARVAKNSQSRVAFEBEYVCPKRNKIC  
 GANIRFPDIIQVTDGSAVPTFYMDFISISGAVSFVVRGEGQVSLTGPANFSYV  
 ALGSRVRELHISLNKTVFLGLLHRKADRIILTKOEARIVADSIDIADVLKEYSY  
 AASALPDEDVEYELLKSGVQVLRGSRLEEIPL"

CDS  
 10777..12432  
 /note="p63; putative heat shock protein 90 homolog"  
 /codon\_start=1  
 /product="63 kDa protein"  
 /protein\_id="AAC40859.1"  
 /db\_xref="GI:3123914"  
 /translation="MSNYSMSLFFKFFYEADMKVYSRIASHSEIKTLPIRLYG  
 GRVYKSFESALPNSPFOELGFLISBRVQSKLGIITVEAAZYLTLYKYPFA  
 ETCSVDKRGESQKYSMEDVAMPWLSLDDNDKLTTCOCSNLSGCELLNPDKRE  
 VLTGSDTLFTAPDRIPILITGLIDYFCSEYANVFYSYKANDNFRFLIANTWPLISDF  
 VFQWKPAPVRLLEFELSAELLLEVPILSIDSQVAVGHILRVESYTSDDPAILDALE  
 DKELAILKSNPRLSTAQWGFVYCYGEFTAQSRVQSGVATPDSVSGGEINIK  
 DYEFEDLQRELPVLSLROFNAGARAEAKFIKPNGNISRPISRLNVRREFYINIK  
 DVFHANRSGLTBEELIINNISVDVRLCERACNTLPKAKRPSKXHKNSINQSRQE  
 RRIIDPLVYLKDTLYEFOHKRAGWSSSTRDLSGRADHAKGSG"

CDS  
 12344..13015  
 /note="p25; coat protein duplicate"

CDS

CDS  
 /codon\_start=1  
 /product="25 kDa diverged coat protein"  
 /protein\_id="AAC40860.1"  
 /db\_xref="GI:3123915"  
 /translation="MSNNTSVPVGGLBALETSGVVLITTRKAVDKFENELKENYSSV  
 DSSLSDSEVVELEKESFKSELSTDEHFWYHIFLIRCAKISTSEKVKVQVGH  
 TYVVDGKYTVLDAWVFNWMSLSLTKYKRVNGJLAPCCACEDLYLVAPIMSRFKTK  
 AVGMKGLPVGKEYIGADPLFGSTSLMSDHDRAVSYVAAKNAVRSATGGERKIVSLY  
 DLGRY"  
 13084..13680  
 /note="p22"  
 /codon\_start=1  
 /evidence=experimental  
 /product="22 kDa coat protein"  
 /protein\_id="AAC40861.1"  
 /db\_xref="GI:3123916"  
 /translation="MELMSDSNLSNLYTTDASSINGVDYDKLISAEVKMLVOKGAPNE  
 GIEVFGLLYALARTTSPKVRADSDVIFSNFGRNVAVTEGDLKIDCGCAPTE  
 RFTNKLRTFRTFTEAVVDPCIAVKKHLPOLNAAALGIPAEBSYLAADFLGTCPKLS  
 ELQOSRKMFASTMAVLTKEGVVNTPVENLRQLGRREVW"  
 13680..14165  
 /note="p19"  
 /codon\_start=1  
 /product="19 kDa protein"  
 /protein\_id="AAC40862.1"  
 /db\_xref="GI:3123917"  
 /translation="MEDYEKSESLILRTNLTMLLVKSDASVELPKLIIQGLRV  
 SGRCVTCNREELTRDPEGNHHTVIRSLIQYSESAFEEFNNSDCVKKFLETGSV  
 FMPFLRSTKRAVRAHRTFPEANNFFGSHGTMEXCLKVLTETSLIDSFCBERN  
 R"

CDS

CDS

3'UTR

ORIGIN

Alignment Scores:  
 Pred. No.: 1,696-95 Length: 15000  
 Score: 991.00 Matches: 198  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0  
 DB: 14

US-09-613-486-15 (1-198) x AF039204 (1-15000)

QY 1 MetGluLeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeu 20  
 Db 13084 ATGAGATTGATGTCGACAGCAACTTAGCAACCTGGTGATATACCGACGCTCTAGTCTA 13143  
 QY 21 AenGlyValAspLysLysLeuLeuSerAlaGluValIGluVSMetLeuValGlnLysGly 40  
 Db 13144 AATGGTGTGACAGAGAGCTTTATCTGCTGAAGTTGAAAAATGTGATGACGAAGAAGG 13203  
 QY 41 AlaProAsnGluGlyIleGluValValPheGlyLeuLeuLeuTyraIleuValaIaArg 60  
 Db 13204 GCTCCTTAAGAGAGGATTAATAATGCTTCCGCTTACTCTTACGACTCGCGGCAAGA 13263  
 QY 61 ThrTh-SerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80  
 Db 13264 ACCACCTCTCTTAAGATTACAGGCCACAGATTAGAGCTTATTTTCAATATGTTCCGA 13323  
 QY 81 GluArgAsnValValThrGluGlyAspLeuLysValLeuAspGlyCysAlaPro 100

Db 13324 GAGAGGAATGTGTAGTAACAGAGGGTGACCTTAAAGAGTACTCCAGCGGTGTGGCCT 13383

Qy 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrArgLysVal 120

Db 13384 CTCACGTAGTTCACCTAATAAAGCTTAGAAGCTGTGGTGTACTTCACTGAGGCTTACGTT 13443

Qy 121 AspheCysIleAlaIleTyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140

Db 13444 GACCTTTGTATCGCGTATTAAGCAAAATTAATCCCACTCAAGCGCGCGGGAATTGAGG 13503

Qy 141 IleProAlaGluAspSerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160

Db 13504 ATTCACGCTGAAGATTGCTACTTACCTGACAGATTCTTGGGTACTTCCCGAAGCTCTCT 13563

Qy 161 GluLeuGlnGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyGly 180

Db 13564 GAATTACAGCAAAAGTAGAGAAATGTTCGCGAGTATGACCTTAAAAACTGAAGGTGGA 13623

Qy 181 ValValaAsnThrProValSerAsnLeuArgGlnLeuGlyArgArgGluValMet 198

Db 13624 GTGGTAATAACACCACTGACAGCAATCTGCTGACCTAGGTAGAGGAAGTTATG 13677

RESULT 3

LOCUS AR138299 15500 bp DNA linear PAT 16-JUN-2001

DEFINITION Sequence 1 from patent US 6197948.

ACCESSION AR138299

VERSION AR138299.1 GI:14479808

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 15500)

AUTHORS Zhu,H.-Y., Ling,K.-S. and Gonales,A.D.

TITLE Grapevine leafroll virus 2 proteins and their uses

JOURNAL Patent: US 6197948-A 1 06-MAR-2001;

FEATURES

source

1. 15500

/organism="Unknown"

/mol\_type="unassigned DNA"

ALIGNMENT Scores:

Pred. No.: 1.76e-95 Length: 15500

Score: 991.00 Matches: 198

Percent Similarity: 100.00% Conservaive: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

Gaps: 0

US-09-613-486-15 (1-198) X AR138299 (1-15500)

Qy 1 MetGluLeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeu 20

Db 13584 ATGAGGTGATGTCCGACAGCAACTTACCACTGGATGATACCGACGCTCTTACGTCTA 13643

Qy 21 AsnGlyValAspLysLysLeuSerAlaGluValGluLysMetLeuValGlnLysGly 40

Db 13644 AATGCTGTGCAAGAGAGCTTTATCTGCTGAGAGTTGTAAGAAAGTTGGTGCAGAAAGG 13703

Qy 41 AlaProAsnGluGlyIleGluValValPheGlyLeuLeuLysTyrAlaLeuAlaAlaArg 60

Db 13704 GCTCTTAACGAGGATATGAAAGTGTGCTGCTACTCTTACGACACTGCGGCAAGA 13763

Qy 61 ThrThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80

Db 13764 ACCACGCTCCCTTAAGGTTAGCGCGAGATTCAGAGCTTAATATTTTCAAAATAGTTTCGA 13823

Qy 81 GluArgAsnValValValThrGluGlyAspLeuLysValLeuAspGlyCysAlaPro 100

Db 13824 GAGAGGAATGTGTGTAGTAACAGAGGATGACCTTAAGAAAGTACTCGACGCGGTGTGGCCT 13883

Qy 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrArgLysVal 120

Db 13884 CTCACGTAGTTCACCTAATAAAGCTTAGAAGCTGTGGTGTACTTCACTGAGGCTTACGTT 13943

Qy 121 AspheCysIleAlaIleTyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140

Db 13944 GACCTTTGTATCGCGTATTAAGCAAAATTAATCCCACTCAAGCGCGCGGGAATTGAGG 14003

Qy 141 IleProAlaGluAspSerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160

Db 14004 ATTCACGCTGAAGATTGCTACTTACCTGACAGATTCTTGGGTACTTCCCGAAGCTCTCT 14063

Qy 161 GluLeuGlnGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGlyGly 180

Db 14064 GAATTACAGCAAAAGTAGAGAAATGTTCGCGAGTATGACCTTAAAAACTGAAGGTGGA 14123

Qy 181 ValValaAsnThrProValSerAsnLeuArgGlnLeuGlyArgArgGluValMet 198

Db 14124 GTGGTAATAACACCACTGACAGCAATCTGCTGACCTAGGTAGAGGAAGTTATG 14177

RESULT 4

LOCUS GLAV4131 8590 bp RNA linear VRL 08-JAN-1998

DEFINITION Grapevine leafroll-associated virus 2 genes encoding RNA polymerase and coat protein, hept70, hept90 gene and ORF2, ORF7 and ORF8.

ACCESSION Y14131

VERSION Y14131.1 GI:2765391

KEYWORDS

SOURCE ORF5; ORF6; ORF7; ORF8; RNA polymerase.

ORGANISM Grapevine leafroll-associated virus 2

REFERENCE 1

AUTHORS Abou-Ghanem,N.

TITLE The nucleotide sequence of the 3' terminal region of grapevine leafroll associated closterovirus 2

JOURNAL Unpublished

REFERENCE 2

AUTHORS Abou-Ghanem,N.

TITLE Direct Submission

JOURNAL Submitted (01-JUN-1997) N. Abou-Ghanem, Universitat degli Studi di Bari, Dipartimento di Protezione delle Pianta e Centro Studio Virus e Virus delle colture mediterranee, via Amendola 165/A, I 70126 Bari, ITALY

REMARK revised by [3]

REFERENCE 3

AUTHORS Abou-Ghanem,N.

TITLE Direct Submission

JOURNAL Submitted (07-JAN-1998) N. Abou-Ghanem, Universitat degli Studi di Bari, Dipartimento di Protezione delle Pianta e Centro Studio Virus e Virus delle colture mediterranee, via Amendola 165/A, I 70126 Bari, ITALY

COMMENT On Jan 10, 1998 this sequence version replaced gi:2369864.

FEATURES

source

1. 8590

Location/Qualifiers

/organism="Grapevine leafroll-associated virus 2"

/mol\_type="genomic RNA"

/db\_xref="taxon:64003"

/lab\_host="N.benthamiana"

1165..2427

/note="ORF1"

/codon\_start=1

/product="RNA polymerase"

/protein\_id="CAA74561.1"

/db\_xref="GI:2369865"

/db\_xref="GOA:O39852"

/db\_xref="SPTRMBL:O39852"

/translation="MRGEMAMNCCRCRDLDAFSLINDVITSIRSGIEQWLGKRP

SRKALKMDVBSPLRIDETCRFKLMVKKDAVKDLSCTTGSAAONIMFRKSIDA

ISPLFNEVKNIMCCLKPNIKRFTENIKRDFASVSNLGDQDYTHICEVDPSKYD

SODAFYKAEVEYMKELGVDEBLAIIMKGEBSLANTIDGQLSTTEIKRSGASNT

WIGNSLVITGLISLYVDVNPALYISGDSLIIFRSEISNYADICTDMGEITFWS

PSVPIYCSKFFVMCGHTEFFVLDPYKLFLKGLAVVEDVSMDSLPETFTFSFKDLTSDFN

CDs

2491. .2651  
/note="ORF2"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAA74562.1"  
/db\_xref="GI:2369863"  
/db\_xref="GOA:O39853"  
/db\_xref="SPTREMBL:O39853"  
/translation="MNVOLEPCLFLNLAVFAVTFILLVPRVKSFRQKHGPAV  
PVRGGGPGSTIV"  
2677. .4476  
/gene="hap70"  
2677. .4476  
/gene="hap70"  
/codon\_start=1  
/product="heat shock protein 70-like"  
/protein\_id="CAA74563.1"  
/db\_xref="GI:2369867"  
/db\_xref="GOA:O39854"  
/db\_xref="SPTREMBL:O39854"  
/translation="WVVFGLDPRGTFSVVCYKXGRVSEKONNSAVITPVLYLFSNS  
NMTGTEHEBESLMSLTKVKGSTFYDLKRWGCDSDNDAYIDRLKPHYSRLYKLTGGG  
LNEVTSIGFPGTGVSEAHLPGLILFLKAVISQEGAFACGTGVCSPANTDSIVQ  
RNFTQVSGISGYOCVWINERINESAALASACNSVEKSNLNLVYFGGTGPDVATISYR  
NNTFVVRASGGDILNGRDVDRAFLTHLFLSLTSLPEDLTLVSNLKSLSLTAELVYI  
TLRGDVRGSDERVKNKILTSMLPEIVARTIKLIESLTLSYAKSMESALVKCDLVYI  
GSSVSLPGIADLTGQSDVRLIRVSDPRAAVAGCALYSCLSGGGLIDCAART  
VAIARSCHOITCAPAGAPVPIPSGSMPLYLARVKNKSGRELAIVEGQVYKCPKMKTC  
GANITFPDITGTDGSIYAPVTPYMDPSISVGASFTVVGPEGKQVSLTGDPAVNSFS  
ALGSSPVERHLISLNKKVFLGLLHRKADRIILFTKQBAIRYADSIDADLYKRYKY  
ASASPPEDEVLILNKGVSQKVLRGSRLEELPL"  
4403. .6058  
/gene="hap90"  
4403. .6058  
/gene="hap90"  
/codon\_start=1  
/product="heat shock protein 90-like"  
/protein\_id="CAA74564.1"  
/db\_xref="GI:2369868"  
/db\_xref="SPTREMBL:O39855"  
/translation="MSNYSWESLFFKPYGEADMKKYLRSIAHSSSEIKLPDRLTNG  
GVYKXSEBSALPNSFPEGLGLTILSRREKQSKMLCGITVEAAVYDLINRKAYKPA  
ETCSPEVKSQGIKSNEDWMTFMRLSNLDVNDKYLAEQCNLSLSCGLIYNPDDKGR  
VALTFEDRIADDTGAANEQVGYLVYAMSLFEQRQKSGSNTSLYERKQITRT  
YLSGDIPLFTAPDRIPGLVGIHIDCEKYNI FYSSYKQSNVNERFELANYPLISDFE  
VYQWYPAADVADLIFELISAAELTLEVPITLIDSGVVGHILRYVESTSPALDALE  
DKLEALIKSNRSLSTALMGVFCFYGFSPFRAQGVORGVVYTPDSVGGFINADKE  
DYKPEPDKICREILPNVSLARQNGARHAHEAFIPNGNISPRLSNLWPEFMYLINI  
DYKPRHANRGLTIEEILLNINISVDYRKLCAERASTLPASAKPESKNKSNIQSRROE  
NRKIDPLVVKOTLIEVFQHKRAGKRSRTIDLSNADAKSGG"  
5970. .6641  
/note="ORF5"  
duplicate gene"  
/codon\_start=1  
/product="coat protein"  
/protein\_id="CAA74565.1"  
/db\_xref="GI:2369869"  
/db\_xref="GOA:O39856"  
/db\_xref="SPTREMBL:O39856"  
/translation="MSSNTPSPVGGLEALETSGVLLYTRKEAVDKFENELKNENYSV  
DSRLSDSEKVEYLEKSKESPSKEIASLTDHEFVYHIFPLIKCAKISTSEKVKYGGSH  
TYVVDKTYTVDIAWAFNMKSLTKYKRVNGVLRACGACEDLVYVAPINSEKRTV  
ANMGKSLPGKEKELGMDFLSGTSKLMSHDRAVSIYAAKNAYDRSAFTGERKIVYSL  
DUGRYV"  
6710. .7306  
/note="ORF6"  
/codon\_start=1  
/product="coat protein"  
/protein\_id="CAA74566.1"  
/db\_xref="GI:2369870"  
/db\_xref="GOA:O39857"

CDs

CDs

CDs

Gene

CDs

Gene

CDs

Gene

CDs

```

/ db_xref="SPTREMBL:O39857"
/translation="MEI MSDSNLSINLVITDASLNGVDKLLSAEYKMLVQKAPNE
GIVEFLLIYALAAFTTS PKVQAPASDVI FENSNGENNVVTEGDLKVDGAPLTL
RFTKLTFFRTFTLEAYVDECIAYKHKLQQLAAALEGI PADSDYLAADFGLGTCPKLS
ELOSRRMPASVYALKTEGGVNVTPPSNLRQLGREVW"
7305..7791
/ note="ORF7"
/ codon_start=1
/ product="hypothetical protein"
/ protein_id="CAA74567.1"
/ db_xref="GI:2369871"
/ db_xref="SPTREMBL:O39858"
/translation="MEDYEPKSPSLTNTLNTMLLVKPSAYELPKLLIGCYLRLY
SGRKEVTCNCRBEILTRDFEGNHTTVRSRIIOYDEBSAEPEPNNSDCVYKFLTGSVY
FWFLRLSRSTGRVLRHRTFFEANNFPSHGQIWECLKQVLTETBSIIDSFCBERN
R"
7793..8410
/ note="ORF8"
/ codon_start=1
/ product="hypothetical protein"
/ protein_id="CAA74568.1"
/ db_xref="GI:2369872"
/ db_xref="SPTREMBL:O39859"
/translation="MRVYVSPYAEADILKRSITDMLNIDSGLSTKECIKAFSITTR
LHKASVQMGVDGLYQRNCAEKRLIDVESNIRLAQPLVREKNAVHCQDEPELVL
AFIRKRYELTGVMREAVRERMSRLTKVLNMSLEMAFVMSPPAMKRAEMLDKFS
PVKIFRDLIDVETINELCAEDDVHYDKVNEDEHNDLELQDEC"

ORIGIN
Alignment Scores:
Pred. No.: 4.57e-95 Length: 8590
Score: 984.00 Matches: 197
Percent Similarity: 99.49% Conservative: 0
Best Local Similarity: 99.49% Mismatches: 1
Query Match: 99.23% Indels: 0
DB: 14 Gaps: 0

US-09-613-486-15 (1-198) x GLAIV4131 (1-8590)
QY 1 MetGIuLeuMetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeu 20
6710 ATGGAGTGGATGATCGGACAGCAACCTTAGCAACCTGGTGTGATTAACCAACCCCTTCATCTCA 6765
QY 21 AANGIYAlAspLyLeuLeuSerAlGIuValGIuLyMetLeuValGlnLyGly 40
6770 AATGGGTGTGACAGCAAGCTTTTATCTGCTGAAGTTGTAAAAATGTTGGTCAGAAAGG 6822
QY 41 AlaProAsnGIuGIYIleGIuValValPheGIYLeuLeuLeuTyrAlaLeuAlaIaArg 60
6830 GCTCCTAACGAGGAGTATGAAAGTGGTGTGGTCTACTCTTACCCACACTCGCGGCAAG 6883
QY 61 ThrThrSerProLyValGIuArgAlaAspSerAspValIlePheSerAsnSerPheGly 80
6890 ACCAGCTCTCCAAAGCTTCAAGCGGCGAGATTCAAGCTTATATTTTCAATAGTTCCGA 6944
Db 81 GIuArgAsnValValIleThrGIuGIYAspLeuLyLyValLeuAspGIYCYAlaPro 100
6950 GAGAGGAATGTGGTGTGATTAACAGAGGTGACCTTAAAGAGGTACTCCACAGCGGTGTGGCT 7003
QY 101 LeuThrArgPheThrAsnLyLeuAsyThrPheGIYArgThrPheThrGIuAlaIYVal 120
7010 CTACTAGGTTACTAAATAACTTGAACGTTGGTGTACTTCTCACTAGAGCTTACGTT 7066
Db 121 AspPheCYAlaIYTYLyHisLyLeuProGIuLeuAsnAlaAlaGIuLeuGIY 140
7070 GACTTTGTATGCGCTAATAGACAAATATACCCCACTCAACCGCGCGGGAATGGGG 7122
QY 141 IleProAlaGIuAspSerTyrLeuAlaAlaAspPheLeuGIYThrCYProLyLeuSer 160
7130 ATTCCAGCTGAAGATTGCTACTTACCTGACAAATTTTCTGGGTACTTCCCGGAAGCTCTCT 7188
QY 161 GIuLeuGIuGInSerArgLyMetPheAlaSerMetTyrAlaLeuLyThrGIuGIYGIY 180

```

Db 7190 GAATACAGCAAGTAGAGAAAATGTTCCGACGATGATGACGCTCTATAAAAATCTGAAAGGTGGA 7249  
Qy 181 ValValAsnThrProValSerAsnLeuArgGlnLeuGlyValArgGlnValMet 198  
Db 7250 GTGGTAATATACCCAGTAGGAAATCTGCGCTACGCTAGGTAGGAAGGAAGTTATG 7303

RESULT 5  
AF314061  
LOCUS  
DEFINITION  
16527 bp ss-RNA linear VRL 06-NOV-2002  
AF314061  
Grapevine rootstock stem lesion associated virus  
methyltransferase/helicase polyprotein, RNA-dependent RNA  
polymerase, p6, HSP70-like protein, HSP90-like protein, coat  
protein duplicate, coat protein, p19, and p24 genes, complete cds.  
AF314061.1 GI:24636914

ACCESSION  
VERSION  
KEYWORDS  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
FEATURES  
SOURCE

1  
16527  
Location/Qualifiers  
/organism="Grapevine rootstock stem lesion associated  
virus"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:167634"  
107..8902  
/note="327 kDa"  
/codon\_start=1  
/product="methyltransferase/helicase polyprotein"  
/protein\_id="AA063466.1"  
/db\_xref="GI:24636915"  
/translation="MSSLALSLPSAAVAGSFGQPAATVAKSFYMTSLPSLQCTTCQ  
EFSFLICLASFQKIKTFLSFLSVREIALFSELTGFCCEFLRGVGHHPFFVAF  
OCGLSVCLGNGIPALINISFPFLDVASLIMGVAVTPFKVIRSKAKNKRAFL  
SIAABAARIRSKVGNKEVRISHOPLAGCLRFSONKNGTELIIGRDVAVARAT  
AAPVRRRGCGRVPLIPLAVSPGVAVRCVACGALCFCKKLASRVAAVDSRAE  
KLUSRRKKLQQQNFNSRVAKKADVTASRDVAEGERTPHHIFGSGFSEPPPIES  
SGMASSVIREPAITVCGASAPAIKEITPRVSKADSPAPVPREKLTIRVAKGGA  
IVISGVHAEVINKIREGLRFRVGDMMFRFSTHGCKIVANVRNTSVWMLCTV  
NGEKSVAPICSDVYVAMLRVSGGKPELVMSRECPDGLCVANHRVYCAFQCP  
RESDPFLGSPSVALKALVLRFGEDALNIGSVYRAFGCDSDSKARSLML  
AGYVGDEVIHSTLSTAVLKGERRADRPKISFGSPCELRVYESVNPAPLRRE  
KYLRYVERNCRGTESADVGSNVRKRDRIKRESHLAGSVNKKRIAGVLRYS  
VGGDDFNHSPITQGHLLVLRKTSVSCVCELSLTPDSRLRPVITKSDVYAMSR  
AAGGRPLVLTWRDKYKNGCYLAHCRVAAVYKLRFPAPVFDGAPPTAKRSRV  
SVLGRSLNLNIGSFTSRGIPEHCDYAAVYKDRFNSAIVAGEKEGEVAASTPRA  
MKOKTIEAVYRLYGSNDLSLKLSEKDLIDPKNDVSLKDRPIRVPEFMSVNTK  
SLTRFYPOPELKFSTSDHPAAASRLNENETVRLCGHSVSDIGCCPLFNHSTK  
ORRHVRCPLVDGKDAORRVARELOYSVRSAGDEKILDEGSNDVGCYPLGVEEHS  
SMMWVGVYDAIVEICSAVTKKKARITVLMVTPGEFLDRECVMSLDCDEVDY  
HADVTYKFGSSCYSHKSLIKDITMTPYVLAGFLSEVMEYELRMGVNPKITKSYV  
CENISCKLIRTRRANSDVYKALRFRPKKRMCLPGYDITYLDSKSPRPVYVNV  
CSAVSKTFEWMVSFTKSKSRVILSGKIHDDVALLDYESSRAVWLAGVSRILA  
SEYLAKNLSHFSGDSFLFATSFVLRREIRNMTNFNELCLQVRAVAFALDSFLD  
LDATMESITDFAECKVAVDLEBLGCLRANAENKIRNLNAGDMAAKLASEVVDIEK  
PAAKAEHRSKSEKKEKHAQKGRGKAGVLEFLDPLVLSASRLFSCTADQRIK  
SVYALIDTAVSFLLDONENETLSPIMVLSKGSYCMFSYLANGSLANRASAADQIRE  
ATLCANVSYSKAKVVTTPSAACAMNENSCGSPGGRYKSMRYTQVLRVLTFFED  
SSYLPLELSAICGATVLRSSGSYSNLNATLIOTISFSELIVRVAVVTHSLT  
TATDGMRSVSFQIISHLVNTGNLAQIARLITVVLPLKCKCGILFMSGDTYSG  
FTKHGVSFSPITNLILYRGLVDLELAAVGVGVDNSKHLPRDVTQARISITVWQIK  
TAVVDGLKCAASKVYSGIVGKTNVGEESRGDDEHYFSCESETPGLRGSRSG

PSILPFSRLVRARILIGRIKRCRSKITESSISLKHHEFTFLCKGSDVGVVEFL  
SEVDADLSSIEEDDETENLIGGFRMADSDLTDMYFPDSEPLASVNGLECAQIR  
GGRRVGMALILNLYVNSLIMARAASKNSTFSLIPSGLDVLRIRNIFSNKESVVE  
LVRYKYNQFLKSNRMRGLGNREHYVNSFGELVPLLNKLSLETIFIVRAVMSKMW  
CLASREKREFAVLCVTLGFPYARLGVGVCDKILSLRRAVDRTVFLKRCDA  
MNGLRCLFHPDNGDSDTSAGLKGGSQTFMLLRVLRPALSLGYACVCSHDL  
FPLCVSLVEEYEMAKIIGLSPSPRSALAPALLOPRFVYSGKSTPLFLFLK  
ISPRVPLIKGFEEFWAGSPLKGRVPLILIAVKNTLRKGLRPRNGSPORAVP  
VIPSNRVYNDGCEKLSERKQVEALREBLLESDKAVYKLNKRNRADSDSEPR  
PMBSRTDVHVSFCGKANSVGTALERAVLVEADKSEKISMAWDLVAKGTVSDBITN  
LSDDEMAIEEYVSEERDDSPKTVRVEVILNLSNPSPEPPIYVDNKGSTIGLNAV  
REFYMOEALFEIHSKLCAYDQLRIVNEFSVAPOENEDAOIYRNGSTIYQGVQ  
RSNLIKDPHDHDFGKISVRRSRNGVYHDLSTLNSLFLAGYFERSFVNTS  
SVDLILKHPDGGCKTTLIDSLFKYRKKEVSMMLTANKSGQVEILKVEKDINSI  
EQQRKDRITPKKSTIITIDALMHRGCDADLIFIDCEFWAGSVAGTACEFPRCHV  
MIFGDSRQIHYIERNEIDKCLYGDLDVFDQCVYNNVSRYCNMDCALSVYDDM  
IATYKGESEKGSNRIEINSVDLVPMDSITYCMLQSEKLEISKFIKRGSKLAV  
LTVHAGQETVAVNLVALKFOEDEPKSIRHITVALSRHDSLTIVNLAAKRGDQIC  
DAIOKANELVNRFRVFPFSFGSVINLDVKVDVDSNRCAASAPLSVINDPLSDVNP  
GTVVDIFGDSLADSPGFBEGAGIYVRDNISSNTTDHCKRV"  
8901..10280  
/note="53 kDa; translated via ribosomal frameshift"  
/codon\_start=1  
/product="RNA-dependent RNA polymerase"  
/protein\_id="AA063467.1"  
/db\_xref="GI:24636916"  
/translation="SAVRSOALPRKRPGLQENTLYSFEARNYNFSTCDRFTSAGFEGA  
MAMNCTIRCFPLETFESSLKDQVITSTRSGIOMEWKERKPSOIKALMDEISPLIDE  
ICRFLKLVKRAKAKLDSKCTKHSAPNOMIHFHRKSNALSPFENKVRIMSCLP  
NIFTEMTNRDPAFVSNSMIGDDVYHIGVDSKIDSDAPKSPREMYRVELY  
DEELAIWMCGERLSIANTLOGQLSFTLENKRSKASATVMIGSLVTLGILSIYDR  
HFEALYVSGDLSLIFSRNKISNYADDIADMGFTKSDSPENDRVYQKAEVALKVGKHTF  
FVPDYPKLFVYKGAIVEDVSMDFLEETFSKEDLTSFSDNDRVYQKAEVALKVGKHTF  
SCNTALASVHICLRNLSFLSKLYPRVKQVYTSVKALFPGSGSLYBSYVTPRG  
HAWYMWDE"  
10373..10543  
/note="6 kDa"  
/codon\_start=1  
/product="p6"  
/protein\_id="AA063468.1"  
/db\_xref="GI:24636917"  
/translation="MDVQLQFECYLINQAVFATFVFLILIRVIKSFRNVAHETPI  
SAVRQGSRTV"  
10552..12351  
/note="65 kDa"  
/codon\_start=1  
/product="HSP70-like protein"  
/protein\_id="AA063469.1"  
/db\_xref="GI:24636918"  
/translation="WVVGLDGRTFTSTVCYKDGKQVYSPKONNSAVIPTLYLFSEET  
NMWYREBESLHNKARKRGSPRLKRWGDSNPSFDYDRLLKPHAVAPYKIGTG  
LNDVTLIGVGTIVSVEHLPELISFKIKNVASCENAFSCCTGVCISVANDVSD  
RNFTQVCVLSGYQCVYWNESAAALACNLLINKSNANLAVDPGGTPTVVISYR  
NNTFVVRASGGDLNIGRDVDRTPFLTHLFSLSLEPDLSDLSINKESLSKDAIV  
TLKSGDDKEDEVKNSLITVLMPLVNRNLILNATIKSYAKMNGRARSVCDDVLI  
IGSSSVPLGEDILSKHRIDRLILKYADPAAVAGCALYSCLSGSQGILVDCIAHT  
IALAGSCDQIICAPAGAPIPSGTPPLIPANNSORQIAPFEGEYVCGKRNKIK  
GSNIRKLDLGATSVASVSPPTPLVDTLSTVSGVLSVNGPEKEXVSLGTPPANSV  
VJGRVVRRLTSLSKRHGLGLLRKADRLRLBEGEAVRYVEIVAVDVLPKERSY  
NASSLPDSDVEPLNGKSVQVLRGRGRLEPL"  
12378..13882  
/note="62 kDa"  
/codon\_start=1  
/product="HSP90-like protein"  
/protein\_id="AA063470.1"  
/db\_xref="GI:24636919"  
/translation="MSNHSWGLSPFKYTYEADWRKLSYTYAHANTITLDPVOYGS  
GRVYKSESAPRSFEGELGLMLSESDVMCKVCGITVEANQITNPAPYPTA  
EKCDPVTKEGQGYTMEVDWNPMSRNLNDIKKILAEQCSJNSCGVILNDDKRT  
VALTFKNELVDSNANIDCRVGDLYVANKLIFNCRSQKSAQGNISLTKCAVIT  
VLENTLTFSPRIPLLTGLIYDFCKEYNIYSTYKRVNDRFFELTNVMPILSDIF  
VRQWVPAIDVRLFLPDLSAELTLEVPLTSDVSGVIGHVLRFEASYADDAIDALE  
EKDALINKSNPILSLAQLMGVPFCYGGFRTAGRRVVRPGVYTPKPNVGGFEINMK

CDS  
 NVEEFPDKIQREVPVSLRQPNGARHAEAFVFKKNIISKEPIRLNIPPEFNTLV  
 DYFRHANSGLSEEEVLININISVVRKLCARACSTLPSARFSGHKSSVPSLRQ  
 KNYRDLALRNSLYEFRHNGRRV"  
 13845..14516  
 /note="25 kDa"  
 /codon\_start=1  
 /product="coat protein duplicate"  
 /protein\_id="AA063471.1"  
 /db\_xref="GI:24636920"

CDS  
 /translation="MSSGTGIVAGSGEALETSGLSPRKNAVDFKFNELNENYSSV  
 DSRSLDQVEKEVLESKESFKSELSAETDEHVVHIIIFELISALISIDSKYKYS  
 KYVBEKTYTQDQWVYPMIKGLTKOYKRVNGLRAFCACEDVLTVAIPASERKTK  
 AIGMKLPYGEKYLADFLISGTSKLSMDRRAVSIIVAKNANVDRAPFGGRKIVSLV  
 DLGRN"  
 14604..15200  
 /note="22 kDa"  
 /codon\_start=1  
 /product="coat protein"  
 /protein\_id="AA063472.1"  
 /db\_xref="GI:24636921"

CDS  
 /translation="WELMSDDNLSCVITTDASLNGVDKILASVITKMLYOKAPSO  
 GLETFGLLVALLAARTSPKVRADSDITFQNTGDKIVTEBDLVKVLBGCAPLN  
 RFTNKLRTGRTFTBAVDFCVAYKHKQPLAAALGIPADSYLAADFLGACPKLS  
 ELQOSRKMFAASYALKTEGGVNVTPVSNLRQLGRREV"  
 15200..15685  
 /note="19 kDa"  
 /codon\_start=1  
 /product="p19"  
 /protein\_id="AA063473.1"  
 /db\_xref="GI:24636922"

CDS  
 /translation="MEICANPEALVLRNTONTLLVKSDEVDNLEPLIGYLYR  
 SRGVDYGNREVDVDFEGAHHTVRSRTVYDESAYKENVNADCVKFKLETGDV  
 FNFPLQSDIKGRARHLRFFPEANNFPSGSHGTMWEYCLKQVLIETESVIESFCERN  
 R"  
 15687..16304  
 /note="24 kDa"  
 /codon\_start=1  
 /product="p24"  
 /protein\_id="AA063474.1"  
 /db\_xref="GI:24636923"

CDS  
 /translation="MRVIVPYEAEDILRSTNMLRNIDSGLADTRRECIRKSTLRD  
 LHKASXOMCUDTDCDDCAEKRLIDVENSLRIAHPLIRKXVATRCXDEPRELV  
 AFTRKIVKLTGVYARERAKIERKTSKVLNKLSTENAFYSPAPAKRAEMLDLRYT  
 PLKTRDLDDVLTALCEDEIHVVVTKODENHDELEDC"

## ORIGIN

Alignment Scores:  
 Prod. No.: 6,44e-88 Length: 16527  
 Score: 921.00 Matches: 180  
 Percent Similarity: 95.96% Conservative: 10  
 Best Local Similarity: 90.91% Mismatches: 8  
 Query Match: 92.94% Indels: 0  
 Gaps: 14

US-09-613-486-15 (1-198) x AFJ14061 (1-16527)

QY 1 MetGluUeuWeteraPserAenUeuSerAenUeuValIleThraSplaSerSerLeu 20  
 Db 14604 ATGGAGTTGATGCCGATGATTAACCTTGAGCGGCTCGCTCAVACCGACGCTTTAATTGG 14663  
 QY 21 AsnGlyValAspLysLysLeuLysSerAlaGluValGluLysMetLeuValGlnLysGly 40  
 Db 14664 AATGGTTCATTAAGAACTGCTCTCGCGAGGCTCAAAAATGCTTGGCAGAAAGGGG 14723  
 QY 41 AlaProGlnGluGlyIleGluValValPheGlyLysLeuLeuTyrAlaLeuValAlaLys 60  
 Db 14724 GCGGCCCGCAAGGATTTGAACAGTCTTGGGCTACTGCTTACACGCTTCGCGACGAGA 14783  
 QY 61 ThrThraSerProLysValGlnArgAlaAspSerAspValIlePheSerAenSerPheGly 80  
 Db 14784 ACCACATCGCCCAAGTTCAGAGGCTGATTCATTAATCTTTCAAAACACTTACGGT 14843  
 QY 81 GluArgSerValValValThrGluLysAspLeuLysValLeuAspGlyCysAlaPro 100

Db 14844 GACACGCTGTGTCGACAGAGAGGATCTCAAGAAAGTCTTGAAGGTCGACACT 14903  
 QY 101 LeuThraRphThraAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrVal 120  
 Db 14904 CTTATATGATTCACCTAACAGAGCTGCGAACTTGGCGCGACGTTTACGAGGCGGTATGTT 14963  
 QY 121 AspPheCysIleAlaTyrLysIleLysLeuProGlnLeuAsnAlaAlaGluLysGly 140  
 Db 14964 GATTTCGCGGTTCACAGCAATAGATCCCTGAGCTCAACGCGCGCGCAATTTGGG 15023  
 QY 141 IleProAlaGluAspSerTyrLeuValAlaAspPheLeuGlyThrCysProLysLeuSer 160  
 Db 15024 ATACACAGAGAGTTGTCCTAGCTGCGGACTTTTAAAGTCTTCCCGCAACTCTCT 15083  
 QY 161 GluLeuGlnGlnSerArgLysMetPheAlaSerMetCysTyrAlaLeuLysThrGluGly 180  
 Db 15084 GAATTACACAGAGATGAAAGATGTTTGCAAGATGATGACCTTTGAAGACTTGAGGGCGGA 15143  
 QY 181 ValValAsnThrProValSerAenLeuArgLysGluValMet 198  
 Db 15144 GTAGTAAATAGCGCCCTTAGTAAATTTGCCCAATTAGTAGAGAGAGATTATG 15197  
 RESULT 6  
 AF190581 15468 bp RNA linear VRL 04-JAN-2000  
 LOCUS Beet yellows virus strain BYV-4, complete genome.  
 DEFINITION AF190581  
 ACCESSION AF190581  
 VERSION AF190581.1 GI:6492367  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Beet yellows virus  
 Beet yellows virus  
 Viruses; ssRNA positive-strand viruses, no DNA stage;  
 Closteroviridae; Closterovirus.  
 REFERENCE  
 1 (bases 1 to 15468)  
 Pareyriyilov, V.V., Haghiwara, Y., and Dolja, V.V.  
 HSB70 homolog functions in cell-to-cell movement of a plant virus  
 Proc. Natl. Acad. Sci. U.S.A. 96 (26), 14771-14776 (1999)  
 MEDLINE  
 PUBMED  
 10611288  
 2 (bases 1 to 15468)  
 Pareyriyilov, V.V., Haghiwara, Y., Alzhanova, D., and Dolja, V.V.  
 Direct Submission  
 AUTHORS  
 TITLE  
 JOURNAL  
 FEATURES  
 SOURCE  
 1. 15468  
 /organism="Beet yellows virus"  
 /mol\_type="genomic RNA"  
 /strain="BYV-4"  
 /db\_xref="taxon:12161"  
 /note="California strain; RNA transcript derived from  
 this cDNA clone causes formation of red necrotic lesions  
 in *Claytonia perfoliata* and systemic infection in  
*Nicotiana benthamiana*"  
 join(108..7997,7999..9381)  
 /note="The polyprotein encompasses domains for papain-like  
 leader proteinase, putative methyltransferase, RNA helicase  
 and RNA-dependent RNA polymerase (RdRp); ORF1a/b;  
 and papain-like leader proteinase is capable of  
 autoproteolytic release from the rest of the polyprotein,  
 and is a replicational enhancer. RdRp is presumably  
 expressed via a +1 ribosomal frameshift"  
 /codon\_start=1  
 /product="p348"  
 /protein\_id="AA014300.1"  
 /db\_xref="GI:6492368"

/translation="WAPLVSAVPSCAPAPAPAPAGASPTVPSPFCVPSYSDISH  
 FRLTSLDFSVRPLFLNARVHLASTDNPSPSLPLGPHATFVLELNGSAPFIS  
 RHIDFVVRPFSVPTEVLVSLSRTSRLPALCDPLYSKGPCEIASFSTPP  
 CLVSNCAVOIPTHAMESIRPTKTLPAGRLLCFCHKYTRKPEPTLIHESGLAKLS  
 ALGYSKNSRPITVKSAGKEYAVERSDRPFERSRROQTPRASHKPKINKAVE  
 PPFPEPRKDKRKRASLPTKDSGFIITGTLRFLSPRKEPRLPFRREVEIVVKK



HAFAVAVSKPRTFRPVATTGAEVNARNTOCSEPRRHPIITLRASATYFPGKMPLOF  
MKKKYVYKSVKSVSCSVTKSPLEALTSTIKKLPKYSNRSERIKYHPIODDEBI  
EYHHLRGKLSVILLIPKGRAYCVVTAATPOYHAALITAGDRPVGELLQVPGGL  
CYLHAAALCCALQKRTFREEDFVGMTPFKVFKRLTEKPSALKHPRGVORS  
LFHCDVASFSSPPYSLPRFTGVVEEAPETISLKRHALESYVERLSHKNDLARS  
VEKDIIDFKBEIKSLSKSKRSVIVPFVMEVAGLTRAPOPLSTHVSYDHPAA  
AGSLLENETIASAKSPBDIGCCPLFIKRGSDYHCRPIYDMKDAORARYRELO  
ARGLVENISREOLVQAORVSCPTLGNCKVSKDYLIMOVYDALINELASMYLKE  
SKAVITMTWMBGLIDREAPALDALCDVAVDRBMVOYKFGSSCYCHKLNINIKI  
MLTAFTPSGMLBSVEMENRMGNVYKTRASNSPIRGKTLRABACTVYKTL  
PRFKTLKTFLSGDIYTLDAKFVSRVFDVYVNSCVSNKTEBWSYIKSSKRYV  
ISGVIRHADVILDKHSECFPAWMLAVGSRRTTERLANLNYEDVACCFITLFL  
FVMSRRAVAEINRSFRMLKMSILSAGIDYFELDIDNLOHLEYSEVERVIAIONG  
EVDGNENRVLTEILAEADRSIAQCSGALSVPOPRGAGSGRSRGSFLVYL  
VEEYGNLEFVSDAVFLVKYKTESDSPIPRVYRMLDLAEASPEVSVSLCAMLV  
BAGAFSSWADRTVSESVKFPVNRVYKRLNEMSAKTLTKKFRPFLSALAKTVY  
RKAVITAEVMEVPESTLDSGEXSAVERGSSVITLLNSGRLDPSGPRSVITEV  
LIDLATKISIEVLKQISVDSTASALVRVSEITLNSRTMGEHCTFKVYLGGF  
LPVPRKCVALCVPGDMATYARFLEYGVDLFLGRSVNSIKYLCVAAGLDSDIVD  
SVVLKSGVAKERVYLGFSKLIKKNLWFRKAVYRTSSITLSEDEYSCDSKRG  
LRGSSRPTLSRLDIFENFKSKLVIENACFAYERIRNMKLYFFPLNSSEER  
RLICAGDFDILSDAFDEDEMLRQAFEOYVSSDSEVTDGKPTVLRVLYANRPT  
ETPNCGKPEFVYKSNYKALYSRLRVLPWDRRLSDSPGLKGNKALAKLKVCY  
ITACEVSOICCLILRLCQCTPCGVRFLPYITYSTCPLSRVAVAVACPLVYRE  
LDGLSDGLTNMGVFRFLVFRALRAUSAVNSALRKITEFTIGNHHPDVAVET  
NEVAREPLSPVDIDVDCDFGSDSESVSDEVASIPRGLHSGRSNSLTLVAVY  
FKLAGRIPLFLRLNFAVAYERELASKRKLTPIGLARLDNFSITSVYLLQEDVS  
VLNAIDVELVLNSGVNVLPVSWRGSLYKLAELVIGSGFASFLGRKCCVSDSC  
SSSNAGCNEMSPVTRKGEKFPVSSSGSTASMYRLALSDIREHVLSTCRVGSDE  
EERKREYTERGHEHTEHDVPIRSHSOPLSGGECSSADEENRANLTHVAKTISE  
RGTETARNRKRTIRAGSERFIATNTSNEQRPITVHSPESRALTNSVPEVYLDL  
ALBELSKLREYDQKYANFNROECDDKEDDFVLAAGVVGSRNSPLKHFNG  
HEFCRSGGLVPGDYDTSKVDITFHTQNFVANSALLSGYSTFTETNSAVLLYE  
APGGKKTTLIKVCEFTFSKNSILITANSRREELIAKVRVILDEGOTPIOTR  
ILITDYSYLMNRGLTKCYLYDECFMVAGAAVACIEFTKDSAILFEDGRQVYIDR  
NELDVAVSLNDRFVDESRYGVSRCMPDCAMSTFYPKATVATTNLSVAGOSM  
QVRLEISVDIVVESSEFVYLMLOSEKKDILSKFSGRSREKPTVILVHAQOETV  
RKNUVTRKEPDDPFRSEHNTIALSRHSEILVYSLSKRDAIOALVYKCOLV  
ARYVPTISFGSTIDISVNSPTSDRSKCKASAEYINSLFESVPGTISVDGDIS  
EMEGTQVESGADVIRDSAPVNSKIDHDQVRSIRSQIIPKRSKLENNYTES  
RNYNFTVCERSFGQEPQAMAVMLERSPLEKAVARSDVAILTEKGVATWMSKES  
PSQRLASLSDLOKPLELEEBITFLMKWDKAYKLDSSCIVKHPAONIMFKAVN  
AIFSPCFDEPKRVITCTNSNIVFTEMTNTSLASIAKEMGSEHYVNIPEFSKD  
KSDAETIKFERTLYSAGPDEDLDVMGOEYTSNATLLDGLSPFVACNORSKGAN  
TWIGNSTITTCILSMAPYTNFRKALFUSGDYALFSSPTIRNSDAMCTGEBETXL  
TSPVYFCSKFTWNGHDVFPVDPPIKYLKGLKSKBVDDEPFLFETFSRLTDL  
VDERVIELTHLVSKTYGESGDYALCAHCHIRSNSSFKLYPRVKVMVAYHGL  
KFVLRFPANCFREKEDTAFGEAYFLTYDET"

CDS

CDS

SYDDANLSSSELLGRVPIKILGRSVEKLDV"  
11311..12972  
/note="64 kDa protein; ORF4"  
/codon\_start=1  
/product="p64"  
/protein\_id="AAFI4303.1"  
/db\_xref="GI:6492371"  
/translation="MTTRFTSPANYWGLFRPFPGGRKMLNSAASVRRPYDRD  
RFSNGEVLRSNRFSDSTGSEVREFFSLTLFTPKTYEVCXKGVAMEQALSQNRRLSD  
VNVSELVNDRKNGCKRNIOTYNEFVKLNGNLAEPLBCHWSLNSCGELNPKD  
TKRPSILFKGDVYESDEAVSSSYDYLSTCLNLYETCNLSNSGKALYDEFLK  
YIVTLVESDPSRLSDNPLVAGVLYMCEBYNTLXSTYIKANESPPCFSLILPL  
SEIFSMNNEQDPRDLLELDTLETLIKTPIINTHDSFTYKRLRLESFEDSN  
EILIKVQVSLTLRNDPELKLQORWGFCYGVRTAQTARVKADEKLPALAGEFT  
INMSGVSEFFDLQKMPKSVRRFCGSLSHAEFSLIKRGVGPPTIRLNPVYKYS  
YLVNDYVHRVAKGLTODELTLSNIEDVAMCCERVALQARRAQRGEKPFQMGK  
VNVESPHARSIRVKNSNLSLNVLMQVAGRRGRNPLARKH"

CDS

CDS

CDS

CDS

ORIGIN

Alignment Scores:  
Pred. No.: 7,31e-21 Length: 15468  
Score: 297.00 Matches: 71  
Percent Similarity: 54.50% Conservative: 32  
Best Local Similarity: 37.57% Mismatches: 84  
Query Match: 29.97% Indels: 2  
DB: 14 Gaps: 2  
US-09-613-486-15 (1-198) x AFI90561 (1-15468)  
QY 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValIleAspIleValys 26  
Db 13671 GCGACTTTGAAAACGTGAGTCTCGCAGCAACGTCGTTTGCACGCTGAGCAGTCGCGAC 13730

QY	27	LeuenseuerragluValGluuysMerLeuValGluuysValaProhangnuuglyile	46
Db	13731	AAACTACGGAGGAGATTTCGAAGACTGTTGAATTGAAAAGGGTTCCGAGAAGTAAAC	13790
QY	47	GluValValPheglyLeuLeuLeuTyraAlaLeuAlaArgThrTherProLyVal	66
Db	13791	GGCTCCGCTTAGACCTTGTGTTGATCCTCGGCGACGATAGTACTTCTATAAAGAT	13850
QY	67	GluArgAlaAspserAspValIlePheSerAsnserPheglyGlu--ArgaenVala	85
Db	13851	AGTGTCCACCGACGCTCTACTTTATCAAAAGCTTCGTCGTAOTGGGAAAGAAATTGTC	13910
QY	86	ValThrGluGlyAspLeuLysValLeuAspGlyCysAlaProLeuThrArgPheThr	105
Db	13911	CTCACTACCGGTGAACGTAGAGTCTCTTCGTGACTCTCAGAACTTTAGAGGAAAGCCT	13970
QY	106	AsnLysLeuArgThrPheglyArgThrPheThrGluAlaTyraValAspPheCysIleAla	125
Db	13971	AACCAATTGCGGTGTTCTGCGGCACTTTCAGAGAAGATTACATTCCTTCGGGAAGAA	14030
QY	126	TyrLysHisLysLeuProGluLeuAsnAlaAlaIleGluLeuGlyIleProAlaGluAsp	145
Db	14031	TACCGAGGAAAGCTCCCTCCATTGCTAGAGCCAAACGTCACGGTCACTGCTGAAGAT	14090
QY	146	SerTyraLeuAlaAlaAspPheLeuGlyThyCysProLysLeuSerGluLeuGlnGlnSer	165
Db	14091	CACATCTTAGTGGCTGATTTTATATTCGACATCAACAGAACTTACTGACCTACACAAAGG	14150
QY	166	ArgLysMetPheAlaSerMetTyraAlaLeuLysThrGluGlyGlyValValaenThrPro	185
Db	14151	CGTGTGCTGTGGCGCGGCAAAAGCCATCACACAAATTCTCG--TCTGAATCACCA	14207
QY	186	ValSerAsnLeuArgGlnLeuGlyArg	194
Db	14208	GTACTAGTTGAAACAGCTGGGTCGT	14234

RESULT 7			
BYU71295	612 bp	RNA	linear
LOCUS	Beet yellows virus coat protein gene, partial cds.		VRL 14-OCT-1996
DEFINITION			
ACCESSION	U71295		
VERSION	U71295.1		
KEYWORDS	GI:1619939		
SOURCE			
ORGANISM	Beet yellows virus		
	Beet yellows virus		
	Viruses; ssRNA positive-strand viruses, no DNA stage;		
	Closteroviridae; Closterovirus.		
REFERENCE	1 (bases 1 to 612)		
AUTHORS	Creamer,R., He,X.H., Yang,C.H. and Grantham,G.		
TITLE	Characterization of the 3'-proximal encoded proteins of beet yellows clostercovirus		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 612)		
AUTHORS	Creamer,R., He,X.H., Yang,C.H. and Grantham,G.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-SEP-1996) Plant Pathology, University of California,		
	Riverside, CA 92521, USA		
FEATURES	Location/Qualifiers		
source	1..612		

CDS

```

/product="coar_protein"
/protein_id="AAB17001.1"
/db_xref="GI:16189404"
/translation="MSASRPSAATFENSLVDQYCHGEODCLRNFEBCILKTK
VPEIDSLGLALVCLYSATGTSIKRSKSVQSTSTFTASGSGSGELFTHTGELRSLDQ
KTLLEGKRNKRLRCRRFQDYISFAEYVGRPLPAEDHYLAADTSTSL
STELDLQGRLLRLAARNATHTESSQSPVTSIKQGRGLANG"

```

ORIGIN

Alignment Scores:

Pred. No.:	1,32e-22	Length:	61
Score:	296.00	Matches:	71
Percent Similarity:	54.50%	Conservative:	32
Best Local Similarity:	37.57%	Mismatches:	84
Query Match:	29.87%	Indels:	2
DB:	14	Gaps:	2

US-09-613-486-15 (1-198) X BYU71295 (1-612)

QY 7 SerxntLeuSerAenValIleThrAspLaseSerLeuAnGlyValAspLyLe 26  
 Db 31 GCGACTTTGAAAACGTGACTGCTGTGACCAACACGTCTTGACGGTGAAGATCGGAC 90  
 QY 27 LeuLeuSerAaGluValGluLyMetLeuValGlnLyGlyAlaProaenGluGlyIle 46  
 Db 91 AAACCTACGGAGAAATTCCGAAGAGTGTTGAAATTGAAAGGGGTTCCGGAAGATAAATC 150  
 QY 47 GluValValPheGlyLeuLeuLeuTyrAlaLeuAlaIaArgThrTrsProLyVal 66  
 Db 151 GGTCTCCGTTAGACTGTTGTTGTACTCTCGTGGACGATAGGTACTCTTAATAAGTT 210  
 QY 67 GluArgAlaSerAspValIlePheSerAnSerPheGlyGlu--ArgAsnValVal 85  
 Db 211 AGGCTCCACCGACGACTCTACTCTTATCATCAAGCTTCGTTGGTATGGAGAGATTGTC 270  
 QY 86 ValThrGluGlyAspLeuLyLyValLeuAspGlyCysAlaProLeuThrArgPheThr 105  
 Db 271 CTCACCTACGGCGAAGCTGAGTCTCTTCTTGACTCTCGAAGAACTTTTAGAGGAGAAAGCT 330  
 QY 106 AsnLyLeuArgThrPheGlyArgThrPheThrGluAlaTyrValaAspPheCysIleAla 125  
 Db 331 AACCAATTCGCGTGTCTCTGCGACATTTTCAGAGAATTACAAATCCTTCGCGAAGCA 390  
 QY 126 TyrIleShiLyLeuProGlnLeuAsnAlaIaIaGluLeuGlyIleProAlaGluAsp 145  
 Db 391 TACCGAGGAACACTGCTCCGATTGCTAGACCAACCGTCACTGCTGCGAAGAT 450  
 QY 146 SerTyrLeuAlaIaAspPheLeuGlyThrCysProLyLeuSerGluLeuGlnIleSer 165  
 Db 451 CACACTACTAGTGTGATTTCATTCGACATCAACAGACATTACCGACCTCAACAAAGT 510  
 QY 166 ArgLyMetPheAlaSerMetTyrAlaLeuLyThrGluGlyValValAsnThrPro 185  
 Db 511 CGTGTGCTGTGGGCGCGAAGAACGCACTCACAGAAATTCTCG--TCTGATTCAACA 567  
 QY 186 ValSerLeuLeuArgGlnLeuGlyArg 194  
 Db 568 GTAACGTATTGAAACACACTGGGTCGT 594

RESULT 8	AF056575	15468 bp	RNA	linear	VRL 02-JUL-1998
LOCUS	AF056575				
DEFINITION	Beet yellows virus Californian isolate, complete genome.				
ACCESSION	AF056575				
VERSION	AF056575.1	GI:3283076			
KEYWORDS					
SOURCE					
ORGANISM	Beet yellows virus				
	Beet yellows virus				
	Viruses; ssRNA positive-strand viruses, no DNA stage;				
	Closteroviridae; Closterovirus.				
REFERENCE	1 (bases 1 to 15468)				
AUTHORS	Peremylov,V.V., Hagiwara,Y. and Dolja,V.V.				
TITLE	Genes required for replication of the 15.5-kilobase RNA genome of a plant closterovirus				
JOURNAL	J. Virol. 72 (7), 5870-5876 (1998)				
MEDLINE	98285746				
PUBMED	9621048				
REFERENCE	2 (bases 1 to 15468)				
AUTHORS	Dolja,V., Peremylov,V. and Hagiwara,Y.				
TITLE	Direct Submission				

JOURNAL	Submitted (31-MAR-1998) Botany and Plant Pathology, Oregon State University, 2082 Cordley Hall, Corvallis, OR 97331, USA
FEATURES	Location/Qualifiers
source	1..15468
	/organism="Beet Yellows virus"
	/mol_type="genomic RNA"
	/isolate="californian"
	/db_xref="taxon:12161"
gene	108..9381
	/gene="ORF1a/b"
CDS	join(108..7983,8045..9381)
	/gene="ORF1a/b"
	/codon_start=1
	/product="fusion protein"
	/protein_id="AAC25115.1"
	/db_xref="GI:3283077"
<hr/>	
mat_peptide	108..1871
	/gene="ORF1a/b"
	/product="papain-like leader peptide"
<hr/>	
mat_peptide	1972..6233
	/function="autoproteolysis"
	/function="replication enhancer"
	/gene="ORF1a/b"
	/product="putative methyltransferase and RNA helicase"
	/note="ORF 1a"
	join(7983,8045..9378)
	/gene="ORF1a/b"
	/product="RNA-dependent RNA polymerase"
	/note="ORF 1b; presumably expressed via a +1 ribosomal frameshift"
	9445..9609
	/function="unknown"
	/note="ORF 2; p6; small hydrophobic protein"
	/codon_start=1
	/product="6.4 kDa protein"
	/protein_id="AAC25116.1"
	/db_xref="GI:3283078"
	/translation="MDCVRSYLLAFGFWICLFCLVFTWVYKQILFNTPPSN
	BARFNRSTV"
	9609..11405
	/function="unknown"
	/note="homolog of HSP70 proteins; ORF 3"
	/codon_start=1
	/product="65-kDa protein"
	/protein_id="AAC25117.1"
	/db_xref="GI:3283079"
	/translation="WVFGDLDFGTFPSVCAVYGEELYLFKQSDAYIPTFVFLHSDT
	OEAFGYDAEVISNDPSVARGFPLDKWIGDEBYGDLKLPKHYTELLKYAOS
	SKETVRLDYGCTGPRONATLPLGLATFVALITLSEAFKCOCTGVCVPAVNCLO
	NRSTESCVNLSGTCVYNNBPSAALSCSKATSPVLYPDGGFVDSVLSAL
	NFVVRASGDMNLGDRIDIAFVHLTKKQQLPVNRKIDISFLKESISKVSLNP
	PVSVENKRVYLVNVSLEAFAPEVETIKIVKVEKYGSRKLEPSVAKLLMV
	GSSSYLPGILSRSSVPEBECVLDDAAAVAGCALSACLRNDSPLLVDCAHN
	LSISKCSISIVCPAGPTLPFGVTVNMTGNSNAVSALLFGDFPKYLNKRIIF
	SGVALANGVSGSNRTVPLTEINVSWGTTFELVPGVTKLGVGNALVDESSY
	QLEERVADHDKENSQKTLIHATLTKPQRQRKLTGDDALFLKLSADYRRDAGFS
	SYDDAVNSSELGLGVITKILGRSVEKLDV"
	11311..12972
	/function="unknown"
	/note="ORF 4"
	/codon_start=1
	/product="64 kDa protein"
	/protein_id="AAC25118.1"
	/db_xref="GI:3283080"
	/translation="WTRRSTPANYWGLPFRPFGGEMKQLMSFAAVRRPYVRD
	YRNSNGVIVSRKVFSDSTGESFRRSLTLTFPKEVCKLGVAAEQLSGNRSLD
	VNVSLELVNKKVGVGKFNQVTEBERVKLNGLAPSLVHEHMSLNSCGELINFKD
	TKRFVLSIFPKGVVESTDEAVSSYLDVLSCLMLYECONLSNSGKALVDFELK
	YVLYLETSLEKYSLSNDPLVAGVYDMKFEVNTLSKTLKNIIBSPDFLSLYPLL
	SEIFSNMWEQAPADVRLFEELDTTELLKPTINTDSTFLYKRLRYLESYEDDSN
	ELIKVYKDSLLTDNBEIKLQAGVGHCVYGFRAQRKVRDABEYLPLAGFT
	LNMSGEFBEIDQKPKSVSVRRCCGSSLSHASEITPKRFGVGPPIRLNYPVYS
	YLVNDYVRHVRKGLVQDELITLISNIEPDAEKCCEVALQRRARORGEKPPQMKG
	VKNESVPHARSIRYKKNDSILNLYMVKDVGARRQRLNPLHRKH"
	12920..13570
	/note="24 kDa protein; ORF 5"
	/codon_start=1
	/product="minor coat protein"
	/protein_id="AAC25119.1"
	/db_xref="GI:3283081"
	/translation="MLAPBGRGDLIHETNTRDAMETFNYSYLAISYVNPKNLNRK
	ETVELDAVIERFSELYITDEDFVKLALPLAANITTSSTNVGAGVEYITGGK
	FLVNDVAVPLIKECMKFKPNPVRFCATFFDPAVILARSILPKFLNITIGKIP
	SGVEFLADPLITATVCLNHERAKVIVLAQRALIDRAVSSVDGKIVSLDRLS"
	13641..14255
	/note="22 kDa protein; ORF 6"
	/codon_start=1
	/product="major coat protein"
	/protein_id="AAC25120.1"
	/db_xref="GI:3283082"
	/translation="MGSAPISATATFENSVLVDQCLHGEDCKLRRNFECKLKG
	VPEDKLGLALGLCYSCATITGTSNKVSVGTSFPIKASFGSGKELFLTHGELSKGYE

mat\_peptide

/gene="ORF1a/b"  
/product="papain-like leader peptide"

mat\_peptide

mat\_peptide

CDS

CDS

CDS

CDS

CDS





```

/c1one="DBYV p36, x19, R9, R3, 36A, R2, 36, p43, 111-7,
112, 115, 1213, 1210, 124, 1313, 1311, 142, 143, 1311,
1311, 142, 143, 156, 1510, 154, 1518, 1520, 169, 1615, 175"
1..107
join(108..7983,8045..9393)
/gene="ORF1a/b"
join(108..7983,8045..9393)
/gene="ORF1a/b"
/codon_start=1
/product="fusion protein of papin-like protease,
methytransferase, RNA helicase"
/protein_id="CA51871.1"
/db_xref="GI:809539"
/db_xref="GOA:008534"
/db_xref="SPTREMBL:Q08534"
/translation="MAFLNYSAYSCAFAPAPAHAGASPIVPSPPCYRYSDDISH
FRLTLDSPSPRLSLNARVHLASNDNPLPLPGFAETPLVLELNGSSAPESIPS
RHIDPVNRPSPSPPEVSVSLRPSRLPALLCPRLYCSGECVCIASGPPPP
CLVNSCYAQIPHAEMESIRPPTKTPAGKFLQFHKRKTTPETLTHESGIALKTS
ALGVTSKPNRPITVKSASGEKYEAIEISKDESRRRQTPRVSRHKPRKINAVE
PFEPBPKDKRRKRALPTDEGFTTGGRLPRLSEBRLPKREVELPVYKX
HAYPAVSKPRTPTPAATGAYVNAQNCSRPNHPLRSGASTYFGFKMLDPRF
MKKEKXYKRSKVSSCSVTKSPLEALAILKNLPQYSNSELKFEYHFGDDDEI
EYHRLGGKISVILLPKBAGCVTAATPQVHALTIARGDRPVGEILQVSGEL
CYLAHALCCALOKRTPREDPFVGMPTKTFVAKLTETLGGSRALKHRYGRQVRS
LFHODVASAPSPPIPLRFITGVBEAPETISLGHKALBETVERVSHKDLKARS
VERKLDIDFOEIKSLSEKRSVTPFYMGEAVQSGULRAIPQNLSPFHSVSDHPAA
AGCILENLETASMAKSFSDIGCCPFLHKRSTDYHVRPIYDMKDQRRSRBLQ
ARGLVENLSREOLEAQAQVAVCPHPLGNCNVSVDLIMQVYDANLIDNIAAMVKE
SKYAVLMTFPGELDERBAFALDALGCDVVDTRDMQYKSGSCYCKVNTKSI
MLPAFTESGNLFSVEMYENMGVNYKILTRASPEIRGKTLRRARCTEYQVCL
PRPKTKTFLSGDIYLDPAKYSRPTDVVNCNVSKTPEMWASYIKSKKAVY
ISGVIRHVDHIDKSECPAVMLAVANRSRTTEFLAKNLMTYTGDSACETITFL
FREMRSRAVAEINRSPFKMKSLISAGLDEPLDNLNQLHLESEVERVSAIQNG
EVCNENRNVLTETIAEADRKSTIAQGLSGALSVTPQPRGLRGSGRSQVFLNL
VEEVGNLPFSGDAVRELIVKFTGDSQPRFVRYVDELAEASPFVSVSLCAML
EAVASFSWADRVSVSEVKTFFVNRVETFLNMSAKLTKEFFFLFSLALAKIV
RKAKVLEAVMEWPESSILSDGSEVAFECSSVITILLNGRLLPGSSPALIEV
LIDLATKISIEVLKQISPADSTASSALRYVLSLTSNRTMGHGIPTKYVLGCF
LPFVRCVVALCVPGMATARFLIEGVDDLFLGRSVNSIKNYLCVAGLVDSIVD
SVVILKSGVAKERYLGFPSKIKNFLNVPKACVVRTSSDLSDEYFCSDESPG
RLRGSSRFTLSRLDIFNFKSKSLVLENACSAVERIERNMKLYFLPANSSEBAR
LIRKACADPYLSDSADDEDEMLRQAEQYSSDDSVYDGKPTVLRSLANVSERFL
ETFCNCPFPVKNYNYKALYSRLRLVLPVDRNLSDPGLKGNKALKAKFKVCV
ITGCEVSOICCLRLTLGWGTACGIVRLFYITYSSTRVSHVVAVACPLVANE
LDGSLDGLTNWGSVPRFLFYALRRALASNSALRKTIETIFGNLHPDPAVLET
NEVABPUSPEVDIDVDCDGSDESVSDEVANSRPPGLHGSRSNELLVSLAVY
FKLARIPRLFLRNFVAVFVERRLSKRLKTFIQLARLPDNPULSVVYLQEDVS
VLNAFIDVELILNCSGVNVLPIVSWRGSLTLAEALVSGFASPLGNRCVSSDC
SSSNAQCNFMSPVRTKGFVPSSSGSSTASMYERLEALSIDIREFVLSGRCGSE
EERKEVTEPGIEHTEDEVPIRSHSQPLSGECSYSEDEENRANLLPHYKYLSE
RQGETARMKRKLTHGVSEFLNINTSNEQRPRIIVHSPESRALNVSVEPYLDEL
ALPFLSCKLREYDOLKVANRQECICDDNEMPLRAGQVSGRNSLPLKHKXG
HEPFRSGGLVPYDGTSDVTIRHTQNPVSNALHSGIYSTPTFTTNLSAVLLXE
APPEGKKTTLIKVCEITFSKNSLITANKSRSEELAKVNRIVLDEGTPQLQIDR
ILITDSYLMNRRGLTCVLYLDCFVYHGAACAIEFTKCDAILFGSROQRYGRG
SELDTAVLSLNRFPVDESRYVGCYRCMDVACMLSTPYPTVATVTVNSAGQSM
QVREIESVDVEYESEYVLTMLQSEKKDLKFGKRSRSEKPTVLTVEHQGETY
KRVLEKTFVDEODPFRSENNHTVALSRHVESLTVSLSRKDDALAOAVKAKQVD
AYRYPTSPGSGTLDSVNPSTSDRSKCSASAPYEYNSFLTSVCGTTSVGGTSS
BEMGTQVPESGADNVIRDSAPVNSKTDHEENLYSVSRMYNTFVCGRSGPQEPQA
MAWVABERSPDLERVAARSDVLAITKGVRTMSKRSEBQDLASSDLCKANLSEB
ITTFKMLKRDARAKLDSCLVHPPPQNIIMPRKAVANALFSPCFQEPKRVITVNS
NIVFTEMTNSVLASIKELGSEHYVNGEIDFOSKDKQDAFIKSPEVTLIASAGF
DEDLIDVMOGEYSNATITLDQLSFSDVQORSGASNTWIGNSITGLISLFTYTN
RFKALFVSGDLSLFSBPIRNSADMCETLGEITKLPSPVFCSKFFVMTGHV
FVPDVKVLTCAIKASKDEVDDEPLFEVFTSPRDLTDLNDERVIELLTHVSKGYE
SGDYVALCAIHCRSNFSSFKLTPYKGVVHYGKLVAKFANCPREKDTAFG
BRFLTLTKLSTLV"
108..7983
/gene="ORF1a/b"
/number=1
7983..8045

```

```

CDS

```

```

/exon

```

```

/gene="ORF1a/b"
/number=1
7983..8045

```

```

CDS

```

```

/exon

```

```

/gene="ORF1a/b"
/number=1
8045..9393
/gene="ORF1a/b"
/number=2
9444..9608
/note="unnamed protein product; putative membrane-binding
protein"
/codon_start=1
/protein_id="CA51864.1"
/db_xref="GI:405627"
/db_xref="SPTREMBL:Q08542"
/translation="MDCVRSYLLARGLICLFLCLVVFIMFVYKQILFRITQSN
EARNHSTVV"
9608..11404
/codon_start=1
/product="heat shock 70-related protein"
/protein_id="CA51865.1"
/db_xref="GI:405628"
/db_xref="GOA:P37092"
/db_xref="SWISS-PROT:P37092"
/translation="MNVVPGIDGRTGTFSSVCAVYGEELYLEKORDSAVITYVFLHSOT
QEVAFGYAEVLSNDLSYRGGFPRDLKRMICDEENRDIYELKRPYKTELLVYROS
SKSVTKLCYGVTPQNAFLPGLATFPAKALITASBAFKQCTGQGVSVYVISAQ
RSFTFSCNLSLGGVNVNPPSAALASAGRIKIGASPVLYVFGGTFVSVYISAL
NNTFVYRSGGDMNLGGHDIDKAEVHELYNKAQLPVNYKIDISLKSLSKVSFLN
PVSQOGRVAVLVNVSLEAVAAPFERTIKIYEVYKCYSSMRLEPNYKALLMY
GSSYLPGLTSLRSIPVDECLVLPDRAAVAGCAVYSALFLGNDSPMLVDCAAH
LSISKYGESIVCVAGSRIPTGVRVTNMGSNASVSAALFEGDPVKRLKRIE
RGDYLVANVGTTGATNRVPTLTINNSVGTITFSLVGPPTGVYKLLIGNAYVDSY
QDGERVADLKHNSDKVLLHALTYQFQCKKLLDGDKALFLRLTMDVRRKARKS
SYDAVANSSELGLRIIPKILRGSREKLDV"
11310..12971
/note="unnamed protein product; putative heat shock
90-related protein"
/codon_start=1
/protein_id="CA51866.1"
/db_xref="GI:405629"
/db_xref="GOA:008541"
/db_xref="SPTREMBL:Q08541"
/translation="VTRFSPTANYWGLFRFPPGGQEWNLNLSAASVSRPYSDD
FRFSGVILSTRGSESTGESFVFEESLLTFPTVEYCKLGVAMETALNGNRISD
YVNSRPNIVDKTVGCKENIQSVTRFYKYNANAEPLVHOMSLNSGCELLNPDD
TKRPSULPKGDLAESTDEDAIVSSSYDIYSHCLNAYETONLSNSGKSLYDEPIK
HYDLTLESDLEKSPSDNPVAGLIDMCEBENTLASYLAKTIBSPDCFSILYPLI
SEVFSMNERRAPDRLLFELDAELLKVTITMHDSITLYKRLRTLESYFDDSN
BLIKVVDLITRNDPRLKLAQRAWGFCYGVFRYQTRKRVGDAEYKLPALGEEV
LIMSVEEFPFELQKQKMSIVRRFCGSLSHBAFVYKRGVGFPPITRLNVVKS
YLVVDYHVRVGLQDELITLINSIEGDAEMECEREVALQAPRAGGGERPFOGMKX
TKNELSPHARSIRYKKNDSLLNIMWDVGARQORRLNPLFRHG"
12919..13569
/codon_start=1
/product="coat protein homolog"
/protein_id="CA51867.1"
/db_xref="GI:405630"
/db_xref="GOA:008538"
/db_xref="SPTREMBL:Q08538"
/translation="MLAPEARGDLIHPEENTRDAMEPFPNSYDLAESEVNPNLNRK
ETDELIGYIRBRKSELYITPDEPYKHLAPLIIAANTTSVKNVYGAAYTIGKK
FLVKQAWFPLIKCKKFNKNPVRFCATFEDAYITASQVYIARLNRITRGKGP
SGYELFGLDFTLAVSCVCLNDEKALIVIQASRAALDRAVSSVDGKIVSLFDGLRS"
13640..14254
/codon_start=1
/product="coat protein"
/protein_id="CA51868.1"
/db_xref="GI:405631"
/db_xref="GOA:008531"
/db_xref="SPTREMBL:Q08531"
/translation="MGSAPISAIATFENVSLADQTLHGDCDKLRKNFECKLKG
VPEDNLGIALGLCYSCATIGTSKNVQPTSTFIKASFGGKELYYLTHGELNLSFLGS

```

exon  
5'UTR  
gene  
CDS  
m1sec\_signal

CDS

ORLLEKPKNLRFCRTFOKDYLSLRKRYGKLPPIAARNRHGLPAEDHYLAADFIIST  
 STELTDLOQSRLLARENATHTEFSSESPTSLKQGRGLGTGR"  
 14251..14793  
 /note="unnamed protein product; orf8"  
 /codon\_start=1  
 /protein\_id="CA51869.1"  
 /db\_xref="GI:405632"  
 /db\_xref="SPTREMBL:008544"  
 /translation="MTSSVLAQTKPLFRYLKGFVYVAFETEESSEAEPLVY  
 LHPDLINRKRKLEASVYDFMSCTLRKSVYSVSESSKSSBDPLSGSKTGS  
 KVNLRKTYTFENGVOYVFMGVLQRCVSSDYLMWENFVGHCHCTLYLCNCELDK  
 SGSELEITRPSKREVLKRM"  
 14768..15301  
 /note="unnamed protein product; orf8"  
 /codon\_start=1  
 /protein\_id="CA51870.1"  
 /db\_xref="GI:405633"  
 /db\_xref="SPTREMBL:008545"  
 /translation="MKFPLKDGRTSRALSRSSLSLRVKEIGTNSQSEISRCVDEFN  
 ELASFNHLITVTEHREMEQHPNOSLTPSPSIGMLKEIRAFLEKRYVTPMHEKTA  
 SDTLNAPLEBYCRITGLAREDALREKRVKSVLPHHSELKFEVTENMFSTELK  
 LNLRLRVISSQILGMAL"  
 15299..15480

3'UTR

## Alignment Scores:

Pred. No.:	1..11e-19	Length:	15480
Score:	286.00	Matches:	71
Percent Similarity:	52.91%	Conservative:	29
Best Local Similarity:	37.57%	Mismatches:	87
Query Match:	28.86%	Indels:	2
		Gaps:	2

US-09-613-486-15 (1-198) x BYVUA (1-15480)

QY 7 SerAsnLeuSerAenLeuValIleThraSpIaSerSerLeuAenglyValaSpIySlyS 26  
 Db 13670 GCGACTTTTGAAGAGCTCTCGACAGCAACCGTTTGACGAGAGAGACTGGCAT 13729  
 QY 27 LeuLeuSerAiaGluValGluIyMeLeuValGluIySglYAlaProAenglyYle 46  
 Db 13730 AAACCTTGAAGAGACTTGAAGAGCTTTGAATTAAGAGGCTCCGAGAGATACCTC 13789  
 QY 47 GluValValPheGlyLeuLeuLeuTyraLeuAlaIaIaArgThrThrSerProLyVal 66  
 Db 13790 GGAATTCGGTTAGACTTTGTTGATTCCTGCTACGATAGGCACTTCCACCAAGAT 13849  
 QY 67 GluArgAlaAspSerAaspValIlePheSerAenSerPhe--GlyGluArgAenVal 85  
 Db 13850 AACGTCACCGAGCTCTACTTCAATCAAGCTTCGTTGGTGGAGAGAGACTGTAC 13909  
 QY 86 ValThrGluGlyAaspLeuLySlyValLeuAaspGlyCySaIaProLeuThrArgPheThr 105  
 Db 13910 CTCACCTACCGGTGAATTCCTTCTGGGGCTCTCAAAAACCTTTGGAGGAGAACT 13969  
 QY 106 AsnLySleuArgThrPheGlyArgThrPheThrGluAlaTyraValaSpPheCySleAla 125  
 Db 13970 AACCAATTCGGGTGTTTCTGCGTACTTTTCAGAAAGCATACATATCTTGGCGAAGAA 14029  
 QY 126 TyrLyShLySleuProGluLeuAenAlaIaIaIaGluLeuGlyIleProAlaGluAasp 145  
 Db 14030 TACCGAGGAATTAACCTCCGATTCGACAGCTAACGTCACGGCTACCGCTGAAGAT 14089  
 QY 146 SerTyrrLeuAlaIaIaAspPheLeuGlyThrCySproLySleuSerGluLeuGlnSer 165  
 Db 14090 CACTACTTACCGCTGACTTATTCGACGTCGACGAGACTCACTGACCTACCAACAAGT 14149  
 QY 166 ArgLySmetPheAlaSerMetTyraLeuLySthrGluGlyValaValaSerThrPro 185  
 Db 14150 CGTGTGCTGTAGCGCGGAGAAAGCCCACTCACAGGAAATTCCTCG--TTCGATTCACCG 14206  
 QY 186 ValSerAenLeuArgGlnLeuGlyArg 194  
 Db 14207 GTAACCACTTGAACAACACTAGGTCGT 14233

RESULT 13

A41914

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

ORIGIN

## Alignment Scores:

Pred. No.:	2..78e-21	Length:	783
Score:	285.00	Matches:	70
Percent Similarity:	53.44%	Conservative:	31
Best Local Similarity:	37.04%	Mismatches:	86
Query Match:	28.76%	Indels:	2
		Gaps:	2

US-09-613-486-15 (1-198) x A41914 (1-783)

QY 7 SerAsnLeuSerAenLeuValIleThraSpIaSerSerLeuAenglyValaSpIySlyS 26  
 Db 119 GCGACTTTTGAAGAGCTCTCGACAGCAACCGTTTGACGAGAGAGACTGGCAT 178  
 QY 27 LeuLeuSerAiaGluValGluIyMeLeuValGluIySglYAlaProAenglyYle 46  
 Db 179 AAACCTTGAAGAGACTTGAAGAGCTTTGAATTAAGAGGCTCCGAGAGATACCTC 238  
 QY 47 GluValValPheGlyLeuLeuLeuTyraLeuAlaIaIaArgThrThrSerProLyVal 66  
 Db 239 GGTCTCGGCTTGAACCTTGTGTTGATTCCTGCTGACGATAGGATCTTCAATAAGTT 298  
 QY 67 GluArgAlaAspSerAaspValIlePheSerAenSerPhe--GlyGluArgAenVal 85  
 Db 299 AGTGTCCAACCGAGCTCTACTTCAATCAAGCTTCGTTGGTGGAGAGAGACTGTTC 358  
 QY 86 ValThrGluGlyAaspLeuLySlyValLeuAaspGlyCySaIaProLeuThrArgPheThr 105  
 Db 359 CTCACCTACCGGTGAATTCCTTCTGGGGCTCTCAAAAACCTTTGGAGAGAACT 418  
 QY 106 AsnLySleuArgThrPheGlyArgThrPheThrGluAlaTyraValaSpPheCySleAla 125  
 Db 419 AACCAATTCGGGTGTTTCTGCGGCTACTTTTCAGAAAGCATACATATCTTGGCGAAGAA 478  
 QY 126 TyrLyShLySleuProGluLeuAenAlaIaIaIaGluLeuGlyIleProAlaGluAasp 145  
 Db 479 TACCGAGGAATTAACCTCCGATTCGACAGCTAACGTCACGGCTACCGCTGAAGAT 538  
 QY 146 SerTyrrLeuAlaIaIaAspPheLeuGlyThrCySproLySleuSerGluLeuGlnSer 165  
 Db 539 CACTACTTACCGCTGACTTATTCGACGTCGACGAGACTCACTGACCTACCAACAAGT 598  
 QY 166 ArgLySmetPheAlaSerMetTyraLeuLySthrGluGlyValaValaSerThrPro 185  
 Db 599 CGTGTGCTGTAGCGCGGAGAAAGCCCACTCACAGGAAATTCCTCG--TTCGATTCACCA 655  
 QY 186 ValSerAenLeuArgGlnLeuGlyArg 194



656 GIMACTAGTTGAAACGCTGGTGT 682

Db

RESULT 14

LOCUS BYU51931

DEFINITION Beet yellow stunt virus helicase gene, partial cds; and RNA replicase, p30, p6, p66, p61, p25, coat protein, p18, and p22 genes, complete cds.

ACCESSION U51931 L20761

VERSION U51931.1 GI:1388128

KEYWORDS

SOURCE

ORGANISM

REFERENCE 1

AUTHORS Karasev, A.V., Nikolaeva, O.V., Koonin, E.V., Gumpf, D.J. and Garsey, S.M.

TITLE Screening of the closterovirus genome by degenerate primer-mediated polymerase chain reaction

JOURNAL J. Gen. Virol. 75 (Pt 6), 1415-1422 (1994)

MEDLINE 94267425

PUBMED 8207405

REFERENCE 2

AUTHORS Karasev, A.V., Nikolaeva, O.V., Mushegian, A.R., Lee, R.F. and Dawson, W.O.

TITLE Organization of the 3'-terminal half of beet yellow stunt virus genome and implications for the evolution of closteroviruses

JOURNAL Virology 221 (1), 199-207 (1996)

MEDLINE 96266425

PUBMED 8661428

REFERENCE 3

AUTHORS Karasev, A.V., Nikolaeva, O.V., Lee, R.F., Wisler, G.C., Duffus, J.E. and Dawson, W.O.

TITLE Characterization of the beet yellow stunt virus coat protein gene phytopathology 88 (10), 1040-1045 (1998)

JOURNAL 4 (bases 1 to 10545)

REFERENCE 4

AUTHORS Karasev, A.V.

TITLE Direct Submission

JOURNAL Submitted (20-MAR-1996) Alexander V. Karasev, University of Florida, Citrus Research and Education Center, 700 Experiment Station Road, Lake Alfred, FL 33850-2299, USA

COMMENT On Jun 25, 1996 this sequence version replaced gi:507953.

FEATURES

source

1..10545

/organism="Beet yellow stunt virus"

/mol\_type="genomic RNA"

/specific\_host="Sonchus oleraceus L."

/db\_xref="taxon:35290"

<1..2133

/function="replication"

/note="ORF1a: similar to helicases of other closteroviruses"

/codon\_start=1

/product="helicase"

/protein\_id="AAC55658.1"

/db\_xref="gi:1388129"

/translation="SSFLCVDMVLRQYCFISNKFYSVGSVYFKAFTTINFKRP TLDSIFEVNHNREYORLKNFLIEGLFINENSGSGSASVGEEDNFEATDMDD RESGFTSTANVDVAGELKIERAQRSGHVSLEKADRICTGYGKFERDEFERK TLVSKFLINTENTANPEPEMILBAHLANSVBEFYLOELTFEIFNKLNRFSLE IVEDPRCTGAGDARLYRSDASLYKDMEEYEVPRFGGLPNDEK SGNLTPANTKFTIANSFTLNSNSYRFTFENDSCIRLYRPAFGGCTHTLTAIFK MEKKRNLIVITANKSSQVETLKIKINDSLRREHTKTLKFKSKARENYPSADSNVY TITSYLMNHGTCQDVLFEDECFVWAGAVTASNTLVRSYFPGDSQIHIERNE YDVAFSFDLRLVAKDRIVGYQSYRCPMDVCGMSKRYENTVAATTNBSBKSLLTI TEINSVDVVASKNTYLFLOJSEKKELEKHLAKGVKATVTVHAGQDITKRVLLV RTKROEDAPSSFNHINVAITRATESITVAVLAARNDNIAAICANALVDKRVLLP HSREGGSALNTIDVEPVYTDNSRCASSAPINDPDEVPGSTSLNFQDTSAESMSQ PFEAGANNVVRDSAKRGSGTDDDEGRV"

<2132..3514

/function="replication"

/note="ORF1b: similar to RdRp's of other closteroviruses;

CDs

10545 bp RNA linear VRL 05-APR-1993

Beet yellow stunt virus helicase gene, partial cds; and RNA replicase, p30, p6, p66, p61, p25, coat protein, p18, and p22 genes, complete cds.

U51931 L20761

GI:1388128

Beet yellow stunt virus

Beet yellow stunt virus

viruses; ssRNA positive-strand viruses, no DNA stage;

Closteroviridae; Closterovirus.

1 (bases 2734 to 5576)

2 (bases 1 to 10545)

Karasev, A.V., Nikolaeva, O.V., Koonin, E.V., Gumpf, D.J. and Garsey, S.M.

Screening of the closterovirus genome by degenerate primer-mediated polymerase chain reaction

J. Gen. Virol. 75 (Pt 6), 1415-1422 (1994)

94267425

8207405

(bases 1 to 10545)

Karasev, A.V., Nikolaeva, O.V., Mushegian, A.R., Lee, R.F. and Dawson, W.O.

Organization of the 3'-terminal half of beet yellow stunt virus genome and implications for the evolution of closteroviruses

Virology 221 (1), 199-207 (1996)

96266425

8661428

(bases 1 to 10545)

Karasev, A.V., Nikolaeva, O.V., Lee, R.F., Wisler, G.C., Duffus, J.E. and Dawson, W.O.

Characterization of the beet yellow stunt virus coat protein gene phytopathology 88 (10), 1040-1045 (1998)

4 (bases 1 to 10545)

Karasev, A.V.

Direct Submission

Submitted (20-MAR-1996) Alexander V. Karasev, University of Florida, Citrus Research and Education Center, 700 Experiment Station Road, Lake Alfred, FL 33850-2299, USA

On Jun 25, 1996 this sequence version replaced gi:507953.

Location/Qualifiers

1..10545

/organism="Beet yellow stunt virus"

/mol\_type="genomic RNA"

/specific\_host="Sonchus oleraceus L."

/db\_xref="taxon:35290"

<1..2133

/function="replication"

/note="ORF1a: similar to helicases of other closteroviruses"

/codon\_start=1

/product="helicase"

/protein\_id="AAC55658.1"

/db\_xref="gi:1388129"

/translation="SSFLCVDMVLRQYCFISNKFYSVGSVYFKAFTTINFKRP TLDSIFEVNHNREYORLKNFLIEGLFINENSGSGSASVGEEDNFEATDMDD RESGFTSTANVDVAGELKIERAQRSGHVSLEKADRICTGYGKFERDEFERK TLVSKFLINTENTANPEPEMILBAHLANSVBEFYLOELTFEIFNKLNRFSLE IVEDPRCTGAGDARLYRSDASLYKDMEEYEVPRFGGLPNDEK SGNLTPANTKFTIANSFTLNSNSYRFTFENDSCIRLYRPAFGGCTHTLTAIFK MEKKRNLIVITANKSSQVETLKIKINDSLRREHTKTLKFKSKARENYPSADSNVY TITSYLMNHGTCQDVLFEDECFVWAGAVTASNTLVRSYFPGDSQIHIERNE YDVAFSFDLRLVAKDRIVGYQSYRCPMDVCGMSKRYENTVAATTNBSBKSLLTI TEINSVDVVASKNTYLFLOJSEKKELEKHLAKGVKATVTVHAGQDITKRVLLV RTKROEDAPSSFNHINVAITRATESITVAVLAARNDNIAAICANALVDKRVLLP HSREGGSALNTIDVEPVYTDNSRCASSAPINDPDEVPGSTSLNFQDTSAESMSQ PFEAGANNVVRDSAKRGSGTDDDEGRV"

<2132..3514

/function="replication"

/note="ORF1b: similar to RdRp's of other closteroviruses;

CDs

10545 bp RNA linear VRL 05-APR-1993

Beet yellow stunt virus helicase gene, partial cds; and RNA replicase, p30, p6, p66, p61, p25, coat protein, p18, and p22 genes, complete cds.

U51931 L20761

GI:1388128

Beet yellow stunt virus

Beet yellow stunt virus

viruses; ssRNA positive-strand viruses, no DNA stage;

Closteroviridae; Closterovirus.

1 (bases 2734 to 5576)

2 (bases 1 to 10545)

Karasev, A.V., Nikolaeva, O.V., Koonin, E.V., Gumpf, D.J. and Garsey, S.M.

Screening of the closterovirus genome by degenerate primer-mediated polymerase chain reaction

J. Gen. Virol. 75 (Pt 6), 1415-1422 (1994)

94267425

8207405

(bases 1 to 10545)

Karasev, A.V., Nikolaeva, O.V., Mushegian, A.R., Lee, R.F. and Dawson, W.O.

Organization of the 3'-terminal half of beet yellow stunt virus genome and implications for the evolution of closteroviruses

Virology 221 (1), 199-207 (1996)

96266425

8661428

(bases 1 to 10545)

Karasev, A.V., Nikolaeva, O.V., Lee, R.F., Wisler, G.C., Duffus, J.E. and Dawson, W.O.

Characterization of the beet yellow stunt virus coat protein gene phytopathology 88 (10), 1040-1045 (1998)

4 (bases 1 to 10545)

Karasev, A.V.

Direct Submission

Submitted (20-MAR-1996) Alexander V. Karasev, University of Florida, Citrus Research and Education Center, 700 Experiment Station Road, Lake Alfred, FL 33850-2299, USA

On Jun 25, 1996 this sequence version replaced gi:507953.

Location/Qualifiers

1..10545

/organism="Beet yellow stunt virus"

/mol\_type="genomic RNA"

/specific\_host="Sonchus oleraceus L."

/db\_xref="taxon:35290"

<1..2133

/function="replication"

/note="ORF1a: similar to helicases of other closteroviruses"

/codon\_start=1

/product="helicase"

/protein\_id="AAC55658.1"

/db\_xref="gi:1388129"

/translation="SSFLCVDMVLRQYCFISNKFYSVGSVYFKAFTTINFKRP TLDSIFEVNHNREYORLKNFLIEGLFINENSGSGSASVGEEDNFEATDMDD RESGFTSTANVDVAGELKIERAQRSGHVSLEKADRICTGYGKFERDEFERK TLVSKFLINTENTANPEPEMILBAHLANSVBEFYLOELTFEIFNKLNRFSLE IVEDPRCTGAGDARLYRSDASLYKDMEEYEVPRFGGLPNDEK SGNLTPANTKFTIANSFTLNSNSYRFTFENDSCIRLYRPAFGGCTHTLTAIFK MEKKRNLIVITANKSSQVETLKIKINDSLRREHTKTLKFKSKARENYPSADSNVY TITSYLMNHGTCQDVLFEDECFVWAGAVTASNTLVRSYFPGDSQIHIERNE YDVAFSFDLRLVAKDRIVGYQSYRCPMDVCGMSKRYENTVAATTNBSBKSLLTI TEINSVDVVASKNTYLFLOJSEKKELEKHLAKGVKATVTVHAGQDITKRVLLV RTKROEDAPSSFNHINVAITRATESITVAVLAARNDNIAAICANALVDKRVLLP HSREGGSALNTIDVEPVYTDNSRCASSAPINDPDEVPGSTSLNFQDTSAESMSQ PFEAGANNVVRDSAKRGSGTDDDEGRV"

<2132..3514

/function="replication"

/note="ORF1b: similar to RdRp's of other closteroviruses;

CDs

10545 bp RNA linear VRL 05-APR-1993

Beet yellow stunt virus helicase gene, partial cds; and RNA replicase, p30, p6, p66, p61, p25, coat protein, p18, and p22 genes, complete cds.

U51931 L20761

GI:1388128

Beet yellow stunt virus

Beet yellow stunt virus

viruses; ssRNA positive-strand viruses, no DNA stage;

Closteroviridae; Closterovirus.

1 (bases 2734 to 5576)

2 (bases 1 to 10545)

Karasev, A.V., Nikolaeva, O.V., Koonin, E.V., Gumpf, D.J. and Garsey, S.M.

Screening of the closterovirus genome by degenerate primer-mediated polymerase chain reaction

J. Gen. Vi



CDS  
 TLVRALEKCLNQAPSWNTTLTNLR"  
 7832..8515  
 /function="unknown"  
 /note="ORF6; similar to analogous BYSV coat proteins of other closteroviruses"  
 /codon\_start=1  
 /product="p25"  
 /protein\_id="AAC55664.1"  
 /db\_xref="GI:1388135"  
 /translation="MPPQGAELVEHNANKSLSIEVSSSETRKVGKFNEDHKTPEKQVNNLNDELREVLGKTELKTLKALDEDIYHVAFFSLASVSTSPVYKGSVSYIDORKYTVNDAMIPOVKILASKNPGLRAFCASLGMYLSVAPDPAFGTBSVGRGAPSGSEYLGADPFTSTCPMSHDAVVALASARNLDSASAQIDKQVSLYDFGKVVYT"  
 8551  
 /note="transcription start site for coat protein subgenomic RNA"  
 8617..9264  
 /note="Cp; ORF7; similar to coat proteins of other closteroviruses"  
 /codon\_start=1  
 /evidence="experimental"  
 /product="coat protein"  
 /protein\_id="AAC55665.1"  
 /db\_xref="GI:1388136"  
 /translation="MAGNDEGSDSSASQTMATKDMIFAPFENFARASATCINGENKKTLPERSVVKQDVTESGIPITIGWTLVALATLSTSSKIDEDTKPLVSAKIDAVNVTIYEDIKNPVNSLTLLKNYKTKLPARTFEEYLRFYROYKHLIPINARAKHGI PADYSIADPVOYTSNLKEHQAVLIEGNNAATASSGTTRESAVNLKILGSSK"  
 9261..9743  
 /function="unknown"  
 /note="ORF8"  
 /codon\_start=1  
 /product="p18"  
 /protein\_id="AAC55666.1"  
 /db\_xref="GI:1388137"  
 /translation="MNVIESIQFTVPVHEKALVLTDDAFIFPHGVQHSPPYVLWYHGEGAMADGVSFTYKELQDVNLGLQYRTKMLRVRTSSSLKILNVRATNAV SIFRCNAPTLIFVTSKGPESMDALVYNGIIGPVSVDPTLLSIVESYGNITHFALF"  
 9746..10306  
 /function="unknown"  
 /note="ORF9; similar to the 20-kDa protein of CTV and to the 21-kDa protein of BYV, two other closteroviruses"  
 /codon\_start=1  
 /product="p22"  
 /protein\_id="AAC55667.1"  
 /db\_xref="GI:1388138"  
 /translation="MKLLSDSYVDSINLANKCLNKLANKSVPLSCIASYNELLIKVMASQSDAIRKNEKNEVLPDSDVMTTDLKMRDRDRLRATINEPFCSSNODVYKMITRNNTNHPNSRERPVFEIQDSFKGVITKISLONLSTYKKEVTITNFKCA LLDVGMGVSVTFEVENYIGENLSEFL"  
 10304..10545  
 /note="similar to 3' UTR's of CTV and BYV"  
 3'UTR  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 8,38e-18 Length: 10545  
 Score: 266.50 Matches: 66  
 Percent Similarity: 48.09% Conservative: 22  
 Best Local Similarity: 36.07% Mismatches: 94  
 Query Match: 26.89% Indels: 1  
 DB: 14 Gaps: 1  
 US-09-613-486-15 (1-198) x BYU51931 (1-10545)  
 QY 11 AenLeuValIleThAspAlaSerIeuAenGlyValAspGlyLeuLeuSerAla 30  
 DB 8704 AATTTCCGTCGCGCGAGCGCACTGTTGATGATGTAACAGAGAGAGTTGTCGAG 8763  
 QY 31 GIUValGluLysMetLeuValGlnLysGlyAlaPheAenGluGlyIleGluValAlaPhe 50  
 DB 8764 GAATTTCTGTGAAGTGAAGACTGAGATGTCACCTGAATCCGCACTTCCCAACACTTG 8823

QY 51 GlyLeuLeuLeuTyrAlaLeuAlaAlaArgThrTherProLysValGlnArgAlaAsp 70  
 DB 8824 GGAATGACGCTGTAGCGCCCTAGCTACTTGAAGACCTTCCCTAAATCGATATGAGAT 8883  
 QY 71 SerAspValIlePheSerAenSerPheGlyGluArgAenValValThrGluGlyAsp 90  
 DB 8884 AAGACACGTTAGTACGCGCTAGATAGATGCGCTCACTGACGATPACTTACGAGAC 8943  
 QY 91 LeuLysValLysLeuAspGlyCysAlaProLeuThrArgPheThrAenLysLeuArgThr 110  
 DB 8944 ATTAAGAACTTCGTAACCTCTTGAAGCTGCTTAAGACTATAGANTTAACCTCGAGTC 9003  
 QY 111 PheGlyArgThrPheThrGluAlaTyrValAspPheCysIleAlaTyrLysHisLysLeu 130  
 DB 9004 TTGCTGCGACCTTTTGAGAGAGACTACTACGCTTGTGAGGCAATCAACACATCTTC 9063  
 QY 131 ProGlnLeuAenAlaAlaAlaGluLeuGlyTyrLeuProLagIleAspSerTyrLeuAlaAla 150  
 DB 9064 CCGAATATCGCAAGACGCAACAAACAGTATTCCTCGCTGATCTCTGATTTTGAAGCT 9123  
 QY 151 AspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSerArgLysMetPheAla 170  
 DB 9124 GACTTCGTCGAACACTGACCACTTATTGAAGACGACGACGACGACGACGACGACGAC 9180  
 QY 171 SerMetTyrAlaLeuLysThrGluGlyGlyValValAenThrProValSerAspLeuArg 190  
 DB 9181 GGTAGAAACGCTGCTACAGACTCTTCGCGACCACTAGGAGGCTCGCGCTAACCTTAAA 9240  
 QY 191 GlnLeuGly 193  
 DB 9241 TATCTTGGG 9249  
 RESULT 15  
 AF501867 672 bp RNA linear VRL 02-APR-2003  
 LOCUS Citrus tristezza virus coat protein gene, complete cds.  
 DEFINITION  
 ACCESSION AF501867  
 VERSION AF501867.1 GI:20454248  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Citrus tristezza virus  
 Citrus tristezza virus  
 Viruses; ssRNA positive-strand viruses, no DNA stage;  
 Closteroviridae; Closterovirus.  
 REFERENCE  
 AUTHORS Roy, A., Ramachandran, P. and Brilansky, R.H.  
 TITLE Grouping and comparison of Indian citrus tristezza virus isolates based on coat protein gene sequences and restriction analysis patterns  
 JOURNAL Arch. Virol. 148 (4), 707-722 (2003)  
 MEDLINE 12548851  
 PUBMED 12664295  
 REFERENCE  
 AUTHORS Roy, A., Ramachandran, P. and Brilansky, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-APR-2002) Plant Pathology, University of Florida, 700 Experiment Station Road, CREC, Lake Alfred, FL 33850, USA  
 FEATURES  
 source  
 1..672  
 /organism="Citrus tristezza virus"  
 /mol\_type="genomic RNA"  
 /db\_xref="taxon:12162"  
 /country="India: Bangalore"  
 1..672  
 /codon\_start=1  
 /product="coat protein"  
 /protein\_id="AA022216.1"  
 /db\_xref="GI:20454249"  
 /translation="MNDPRTKLIKNNKKEKEDGDVVAEESFGSNLHIDPTLIANDVROLGQNALNRDLFTLKSKYRPLPDEDLPHIAMLVLAAYKSSLSGDDPTG IYTRGVGVNDLSDKMTDVAIVNSKIGIKRITAKRWGATNTALVIAFORORNSISG GRPLDAGIPAGHYICADFLTAGLTLDECAVYIQAKQLKRGADDEVVTVNVLQSG KFNTR"  
 CDS



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - nucleic search, using frame\_plus.p2n model

Run on: April 7, 2004, 01:14:14 ; Search time 306 Seconds  
(without alignments)  
2748.834 Million cell updates/sec

Title: US-09-613-486-15  
Perfect score: 991  
Sequence: 1 MELMSDSNLSNLVITDASSL.....GGVNTNPVSNALQGRREVM 198

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=x1h  
-Q=/cgn2\_1/USPRO.spool/US0961386/runcat.05042004.15120.28252/app\_query.fasta\_1.391  
-DB=Geneseq.29Jana04 -QPM=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOFC1=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=numat40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US0961386 @CGN 1.1 352 @runcat.05042004.15120.28252 -NCPU=6 -ICPU=3  
-NO.MMAP -IARGCOTRY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOC  
-DRV TIMEOUT=120 -MAXN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq.29Jana04:

1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991	100.0	597	2	AAV08870 Grapevine
2	991	100.0	15500	2	AAV08874 Grapevine
3	286	28.9	688	2	AAQ87853 Sugar beee
4	285	28.8	783	2	AAAT42975 BYV coat
5	89.5	9.0	1500	6	ABQ76539 C. albica
6	86	8.7	1197	3	AAQ45864 Arabidops
7	86	8.7	1199	3	AAAC39012 Arabidops
8	86	8.7	34980	5	AAAF86431 Pyrococcu

9	85	8.6	4467	2	AAT68648	AAT68648 PSR0800 f
10	85	8.6	5311	9	ADD29639	Add29639 Mouse tum
11	84.5	8.5	4766	4	AAA91259	AAA91259 GILAV-5 n
12	83.5	8.4	1266	7	ACA29885	ACA29885 Prokaryot
13	83.5	8.4	110000	6	ABA92787-4	Continuation (5 of
14	82.5	8.3	1773	3	ACA47601	ACA47601 Prokaryot
15	81.5	8.2	1209	3	AAA72313	AAA72313 Erysipelo
16	81.5	8.2	1209	3	AAA72318	AAA72318 Erysipelo
17	81.5	8.2	2040	3	AAA50205	AAA50205 Erysipelo
18	81.5	8.2	2442	7	ACA23154	ACA23154 Prokaryot
19	81.5	8.2	2814	3	AAA72317	AAA72317 Erysipelo
20	81.5	8.2	110000	2	AAK20248-02	Continuation (3 of
21	81.5	8.2	349980	6	ABQ81849	ABQ81849 Bifidobac
22	80.5	8.1	10709	4	AAE28522	AAE28522 Genomic f
23	80	8.1	2136	6	AAE70626	AAE70626 Streptoco
24	80	8.1	2136	7	ACA50630	ACA50630 Prokaryot
25	80	8.1	2187	6	AAE65303	AAE65303 Streptoco
26	80	8.1	110000	6	ABA90521-09	Continuation (10 o
27	79	8.0	1353	7	ACA33088	ACA33088 Prokaryot
28	79	8.0	1356	4	AAE52987	AAE52987 Enterococ
29	79	8.0	1434	2	AAE72221	AAE72221 Grapevine
30	79	8.0	1434	7	ACA43662	ACA43662 Prokaryot
31	79	8.0	2572	4	ABL02037	ABL02037 Drosophil
32	79	8.0	4827	4	ABL02036	ABL02036 Drosophil
33	79	8.0	11597	2	AAE13159	AAE13159 Enterococ
34	79	8.0	11597	6	ABE98954	ABE98954 Enterococ
35	79	8.0	17919	3	AAZ49200	AAZ49200 Grapevine
36	79	8.0	61313	4	AAE59545	AAE59545 Propionib
37	79	8.0	61313	7	ACF64474	ACF64474 Propionib
38	79	8.0	110000	2	AAV30458-3	Continuation (4 of
39	79	8.0	110000	2	AAV30459-3	Continuation (4 of
40	78.5	7.9	2282	4	AAH42270	AAH42270 Nucleoid
41	78.5	7.9	110000	6	ABA90521-10	Continuation (11 o
42	77.5	7.8	665	6	ABK78469	ABK78469 Bacillus
43	77.5	7.8	14494	4	ABJ17446	ABJ17446 Drosophil
44	77	7.8	912	4	AAH02200	AAH02200 Aspergill
45	77	7.8	929	4	AAH01693	AAH01693 Aspergill

## ALIGNMENTS

RESULT 1	AAV08870	standard, CDNA, 597 BP.
ID	AAV08870	
XX	AAV08870;	
AC	27-AUG-2003	(revised)
DT	29-MAR-1999	(first entry)
XX		
DE	Grapevine leafroll virus type 2 coat protein ORF6 product.	
XX		
KW	GRLAV-2; closterovirus; grape; tobacco; transgenic plant;	
KW	disease resistance; virus resistance; beet yellows virus;	
KW	coat protein; ss.	
XX		
OS	Grapevine leafroll virus.	
XX		
PN	WO9653055-A1	
XX		
PD	26-NOV-1998.	
PF	20-MAY-1998;	98WC-US010313.
XX		
PR	20-MAY-1997;	97US-0047194P.
XX		
PA	(CORR) CORNELL RES FOUND INC.	
XX		
PI	Zhu H, Ling K, Goncalves D;	
XX		
DR	WPI: 1999-045307/04.	
XX		
DR	P-RSDB; AAW73482.	
XX		



PI Zhu H, Ling K, Gonsalves D;  
XX  
XX WPI: 1999-045307/04.  
DR P-PSDB: AAM73476, AAM73477, AAM73478, AAM73479, AAM73480, AAM73481,  
DR AAM73482, AAM73483, AAM73484.  
XX  
XX Grapevine leafroll virus (type 2) proteins and polypeptides - and  
PT encoding DN, useful e.g. to impart grapevine leafroll resistance to  
PT grape and tobacco plants and detect grapevine leafroll virus.  
XX  
XX Disclosure: Page 12-20; 151pp; English.

A total of 15,500 bp of the RNA genome of grapevine leafroll virus type (GRLV-2) has been sequenced and cloned from GRLV-2 isolated from infected Vitis vinifera cv. Pinot Noir. About 85% of the total RNA genome was revealed from 2 different clones. The sequence in the coat protein gene region was determined and confirmed from several overlapping clones. The genomic organisation of GRLV-2 includes 9 open reading frames (see also AAV08864-7) encoding a polyprotein (see AAV73476) having papain-like protease, methyltransferase and helicase motifs, an RNA-dependent RNA polymerase, heat shock proteins, coat proteins and other proteins of unknown function (see AAV73476-84). These proteins can be used to produce antibodies, useful to detect GRLV in samples e.g. by ELISA (claimed). The nucleic acid molecules can be used to produce probes and primers for such detection, and to transform host cells (especially Agrobacterium strains, Agrobacterium tumefaciens, grape, citrus, beet or tobacco cells) and produce transgenic plants (claimed). In particular, they can be used to impart GRLV-2 resistance to Vitis scion or rootstock cultivars or Nicotiana (claimed). Because extensive similarity exists between hsp70-related sequence regions of GRLV-2 and other closteroviruses, the DNA may also be used to impart beet yellows virus resistance to beet cultivars or tristeca virus resistance to citrus scion cultivar/rootstock cultivars (claimed). (Updated on 27-Aug-2003 to correct OS field.)

```

US-09-613-486-15 (1-198) x AAO87853 (1-688)
QY 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLys 26
DB 104 GCACCTTTGAAAGCGTAAGTCTCGACAGACCAACCTGTTGACGAGAAAGCTGCGAT 163
QY 27 LeuLeuSerLagluValGluLysMetLeuValGluLysGlyValAProLengluGlyLe 46
DB 164 AAACCTTAGGAAGAACTTCGAAAGGTGTTGAAATTAAGAGGGGTTCGGAAGTAACCTC 223
QY 47 GluValValPheGlyLeuLeuLeuTyralaLeuAlaIleArgThrHisSerProLysVal 66
DB 224 GGAAATCCGCTAGACCTTGTGTTGATTCCTGTGCTACGATAGGACCTTCCAAACAAAGTT 283
QY 67 GluArgAlaAspSerAspValIlePheSerAsnSerPhe---GlyGluArgAsnValVal 85
DB 284 AACGTCACACCGACGTCTACCTTCATCAAAAGCTTCGTGGTGAGGAGAAAGAACTGATC 343
QY 86 ValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaProLeuThrArgPheThr 105
DB 344 CTCACCTACCGGTGAATGATTCCTTCTGGGCTCAAAAACCTTTGGAGGAAACCT 403
QY 106 AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyralValAspPheCysIleAla 125
DB 404 AACAAATTCGGGTGTTCTGCGCTACCTTTCAGAAAGACATCAATCTTCGCGCAAGAA 463
QY 126 TyrlshLysLysLeuProGluLeuAsnAlaIleGluLeuGlyLeProAlaGluAsp 145
DB 464 TACCGAGGAAATTAACCTCCAGTCCAGACCTACCGTCAACCGCTGATGAGAT 523
QY 146 SerTyrlLeuAlaIleAspPheLeuGlyThrCysProLysLeuSerGluLeuGluGlnSer 165
DB 524 CACTACTTACCGCTGATTCATATCGACGCGAACCTCACTGACCTACCAACAAAGT 583
QY 166 ArgLysMetPheAlaSerMetTyrlAlaLeuLysThrGluGlyValValAsnThrPro 185
DB 584 CGTCTGCTGTAAGCGCGCAAAACGCTCAACAGGAATTCCTG---TCTGAATCAACG 640
QY 186 ValSerAsnLeuArgGlnLeuGlyArg 194
DB 641 GTAACCACTTGAACAAACACTAGTGT 667

RESULT 4
AAT42975.
ID AAT42975 standard; cDNA, 783 BP.
XX
AC AAT42975;
XX
DT 16-OCT-2003 (revised)
XX 14-JAN-1997 (first entry)
XX
DE BYV coat protein cDNA.
XX
KM Tobacco mosaic virus; coat protein; beet yellow virus; BYV;
KM beet western yellow virus; BWV; beet necrotic yellow vein virus; BNWV;
KM tobacco etiolation virus; sugar beet; fruit; mango; apple; pear; banana;
KM field crop; sunflower; wheat; barley; maize; vegetable; potato; carrot;
KM cabbage; onion; ss.
XX
OS Beet yellows virus.
XX
XX
XX Key Location/Qualifiers
XX sig_peptide 13..78
XX FT /tag= a
XX FT /note= "Omega sequence from TMV"
XX CDS 89..703
XX /tag= b
XX /product= "BYV_coat_protein"
XX
XX MO9428147-A1.
XX
XX 08-DEC-1994.

```

---

```

XX 01-JUN-1994; 94MO-EP001786.
PF 02-JUN-1993; 93GB-00011332.
XX
XX (SANO ) SANDOZ LTD.
PA (SANO ) SANDOZ PATENT GMBH.
PA (SANO ) SANDOZ-ERFINDUNGEN VERM GES MBH.
PI Boisen K, Brunstedt J;
XX WPI, 1995-022818/03.
XX
XX Recombinant DNA sequence comprising at least two coat protein genes -
XX used to combat viruses in e.g. sugar beet.
XX
XX Disclosure; Page 11-12; 23pp; English.
XX
XX The sequences given in AAT42975-77 represent cDNA sequences encoding coat
XX proteins from beet yellow virus (BYV), beet western yellow virus (BWV)
XX and beet necrotic yellow vein virus (BNWV) respectively. These sequences
XX are used in the recombinant DNA sequence of the invention. This comprises
XX nucleotide sequences encoding the coat proteins of at least one
XX tobacco etiolation virus and at least one closterovirus. This recombinant DNA may be
XX used to transform sugar beet for combating viral infection. Other plants
XX may also be transformed e.g. fruit such as mangoes, apples, pears, and
XX bananas, and field crops such as sunflower, wheat, barley, maize, and
XX vegetables such as potatoes, carrots, cabbage and onion. (Updated on 16-
XX OCT-2003 to standardise OS field)

SQ Sequence 783 BP; 223 A; 178 C; 171 G; 211 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,896-26 Length: 783
Score: 285.00 Matches: 70
Percent Similarity: 53.44% Conservative: 31
Best Local Similarity: 37.04% Mismatches: 86
Query Match: 28.76% Indels: 2
DB: Gaps: 2

US-09-613-486-15 (1-198) x AAT42975 (1-783)
QY 7 SerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLys 26
DB 119 GCACCTTTGAAAGCGTAAGTCTCGACAGACCAACCTGTTGACGAGTGAAGATCGCAG 178
QY 27 LeuLeuSerLagluValGluLysMetLeuValGluLysGlyValAProLengluGlyLe 46
DB 179 AAACCTTAGGAAGAACTTCGAAAGGTGTTGAAATTAAGAGGGGTTCGGAAGAAACCTC 238
QY 47 GluValValPheGlyLeuLeuLeuTyralaLeuAlaIleArgThrHisSerProLysVal 66
DB 239 GGATTCGCGTGAACCTTGTGTTGATTCCTGTGCGACATAGGATCACTTCAATAAGAT 298
QY 67 GluArgAlaAspSerAspValIlePheSerAsnSerPhe---GlyGluArgAsnValVal 85
DB 299 AGTGTCACACCGACGCTACCTTTCATCAAAAGCTTCGTTGGTGAGGAGAAATTTGTC 358
QY 86 ValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaProLeuThrArgPheThr 105
DB 359 CTCACCTACCGGTGAATGAGTCTTTCGACACTCCAGAAACCTTTGGAGAGAAAGCCT 418
QY 106 AsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyralValAspPheCysIleAla 125
DB 419 AACCAAGTTCGTTGTTCTGCGCACTTTCAGAAAGATCAATATCTTCGGAAGAA 478
QY 126 TyrlshLysLysLeuProGluLeuAsnAlaIleGluLeuGlyLeProAlaGluAsp 145
DB 479 TACCGAGGAAGATGCTGCTCGATGCTAGAGCCAAACCTCAACGCTACCTGTAAGAT 538
QY 146 SerTyrlLeuAlaIleAspPheLeuGlyThrCysProLysLeuSerGluLeuGluGlnSer 165
DB 539 CACTACTTACCTGATTCATATCGACATCAACAAACCTTACTGATCAACAAAGT 598

```



Fri Apr 9 15:30:05 2004

us-09-613-486-15.p2n.rng

Page 6

```
XX 25-FEB-2000; 2000EP-00301439.
PF 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123160P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129844P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130444P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-013256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135622P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137503P.
PR 07-JUN-1999; 99US-0137722P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138547P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142255P.
PR 06-JUL-1999; 99US-0142390P.

PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 25-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148655P.
PR 13-AUG-1999; 99US-0148684P.
PR 15-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0152363P.
PR 07-SEP-1999; 99US-0153070P.
PR 10-SEP-1999; 99US-0153758P.
PR 13-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
```



PR 24-SEP-1999; 99US-0155659P.  
 PR 26-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159283P.  
 PR 13-OCT-1999; 99US-0159284P.  
 PR 13-OCT-1999; 99US-0159285P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 21-OCT-1999; 99US-0160816P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:  
 Pred. No.: 1.16 Length: 1197  
 Score: 86.00 Matches: 49  
 Percent Similarity: 38.92% Conservative: 23  
 Best Local Similarity: 26.49% Mismatches: 93  
 Query Match: 8.68% Indels: 20  
 DB: 3 Gaps: 9

US-09-613-486-15 (1-198) x AAC45964 (1-1197)  
 QY 7 SeranSeuSerAenValIleThrAspAlaSerSerLeuAnGlyVal-----Asp 24  
 Db 427 TCGGAACAACCAACATGTCACAGGTACATGCTGCTCAAGGATCGCGAGAC 466  
 QY 25 LysLysLeuLeuSerAlaGluValGluLysMetLeuValGlnLysGlyAlaProAsnGlu 44  
 Db 487 AAGAAAGCAGATGAGATCAATAACATGCCCTGCTGCTGAGTCGAGCGGCTTAACGAT 546  
 QY 45 GlyIle-----GluValValPheGlyLeuLeuLeuValAlaLeuAlaArg 60  
 Db 547 TTCAATTTGAACATTAACGAATTCCTCCAGGGCTCCAGATTCCTTCATTCGT 606  
 QY 61 ThrThrSerProLysValGluThrAlaAspSerAspValIlePheSerAsnSerPheGly 80  
 Db 607 TACCAAGACTTATCTTACAGAGGCTTGAATAATTCGCGGAGCCTTACAGTTAGT 666  
 QY 81 GluThrAsnValValValThrGluGlyAspLeuLysLysValLeuAspGlyCysAlaPro 100  
 Db 667 GTCCGAGATGTTTGGT-----GGAGCTTACCGCCCAAG-----GGTGTTCACG 714  
 QY 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaValVal 120  
 Db 715 ATC--CATATGACTGCTAAATTCGCAACATTTTATAGATTCCTTGGAAACACCATAC 771  
 QY 121 AspPheCysIleAlaValTyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140

Db 772 AAAGACTCTGTTTATACATAGAGAACTTCGATCTTCCCAATGCAAGCATCT 831  
 QY 141 IleProAlaGluAspSerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160  
 Db 832 CTTCACAGA---AGCAAAATCTCTTACCCCAATGCTCTATAT-----CTATATGATGAG 862  
 QY 161 GluLeuGlnGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLys--ThrGluGly 179  
 Db 883 ATGATCCCAAAACCCCTACGAAA-----TACGGTTCAAAGAGACGAAAGAGA 927  
 QY 180 GlyValValAsnThr 184  
 Db 928 GGATGTTGTGMAACA 942  
 RESULT 7  
 AAC39012  
 ID AAC39012 standard; DNA; 1199 BP.  
 AC AAC39012;  
 XX 17-OCT-2000 (first entry)  
 DT  
 XX Arabidopsis thaliana DNA fragment SEQ ID NO: 23034.  
 DE  
 XX Hybridization assay; genetic mapping; gene expression control;  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KM Promoter; termination sequence; ss.  
 XX Arabidopsis thaliana.  
 OS  
 PN EP1033405-A2.  
 PD  
 XX 06-SEP-2000.  
 PF  
 XX 25-FEB-2000; 200DEP-00301439.  
 XX  
 XX 25-FEB-1999; 99US-0121825P.  
 XX 05-MAR-1999; 99US-0123180P.  
 XX 09-MAR-1999; 99US-0123548P.  
 XX 23-MAR-1999; 99US-0125788P.  
 XX 25-MAR-1999; 99US-0126264P.  
 XX 29-MAR-1999; 99US-0126785P.  
 XX 01-APR-1999; 99US-0127462P.  
 XX 06-APR-1999; 99US-0128234P.  
 XX 08-APR-1999; 99US-0128714P.  
 XX 16-APR-1999; 99US-0128845P.  
 XX 19-APR-1999; 99US-0130077P.  
 XX 21-APR-1999; 99US-0130449P.  
 XX 23-APR-1999; 99US-0130510P.  
 XX 28-APR-1999; 99US-0130891P.  
 XX 30-APR-1999; 99US-0132048P.  
 XX 30-APR-1999; 99US-0132407P.  
 XX 04-MAY-1999; 99US-0132484P.  
 XX 05-MAY-1999; 99US-0132485P.  
 XX 06-MAY-1999; 99US-0132486P.  
 XX 06-MAY-1999; 99US-0132487P.  
 XX 07-MAY-1999; 99US-0132863P.  
 XX 11-MAY-1999; 99US-0134256P.  
 XX 14-MAY-1999; 99US-0134218P.  
 XX 14-MAY-1999; 99US-0134219P.  
 XX 14-MAY-1999; 99US-0134221P.  
 XX 14-MAY-1999; 99US-0134370P.  
 XX 18-MAY-1999; 99US-0134768P.  
 XX 19-MAY-1999; 99US-0134941P.  
 XX 20-MAY-1999; 99US-0135124P.  
 XX 21-MAY-1999; 99US-0135353P.  
 XX 24-MAY-1999; 99US-0135629P.  
 XX 27-MAY-1999; 99US-0136021P.  
 XX 27-MAY-1999; 99US-0136392P.  
 XX 28-MAY-1999; 99US-0136782P.  
 XX 01-JUN-1999; 99US-0137222P.

PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137562P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 18-JUN-1999; 99US-0139817P.  
PR 21-JUN-1999; 99US-0139899P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0145442P.  
PR 14-JUL-1999; 99US-0145624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145087P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.

PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148177P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149425P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150894P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151303P.  
PR 01-SEP-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 24-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0156559P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157655P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158832P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 14-OCT-1999; 99US-0159781P.  
PR 18-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161466P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:

Pred. No.: 1.16 Length: 1199  
 Score: 86.00 Matches: 49  
 Percent Similarity: 38.92% Conservative: 23  
 Best Local Similarity: 26.49% Mismatches: 93  
 Query Match: 8.68% Indels: 20  
 DB: 3 Gaps: 9

US-09-613-486-15 (1-198) x AAC39012 (1-1199)

QY 7 SerAsnLeuSerLeuVal11eThrAspAlaSerSerLeuAsnGlyVal1-----Asp 24  
 Db 429 TCGGAACACCAACCAAGTTCAGAGTTCATGCTGCTCAAGAGTATCCTTCACTTCTGCT 488  
 QY 25 LysLeuLeuLeuSerAlaGluValGluLysMetLeuValGluLysGlyAlaProAsnGlu 44  
 Db 489 AAGAAAGCGATGAGATCATAAACATGCTTCTGCTGCTGCTGAGTGCAGGCGCTTAACAT 548  
 QY 45 GlyLe-----GluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAlaArg 60  
 Db 549 TTCATTTTGAACATATACGAATTCCTCAAGGCGCTCGAGTATCCTTCACTTCTGCT 608  
 QY 61 ThrTherProLysValGlnArgAlaAspSerAspVal11ePheSerAsnSerPheGly 80  
 Db 609 TACCAAGACTTATTCCTTAAGAGCTTGAATAATTCCTGCGGAGCTTACAGTTAGGT 668  
 QY 81 GluArgAsnValVal11eThrGluGlyAspLeuLysValLeuAspGlyCysAlaPro 100  
 Db 669 GTCCGGAATGTTTGGTT-----GGAGTTTACCGCGCATG-----GGGTGTTACCG 716  
 QY 101 LeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrVal 120  
 Db 717 ATC---CATATGATGCTGCTAAATTCGACACATTTTATGATTCCTGTTGAACACCATAC 773  
 QY 121 AspPheCysIleAlaTyrLysSHSLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140  
 Db 774 AAAGACTCTGTTTATACATGAGAACTTCAGATCTCTTGGCCCAATCGAAGCATCT 833  
 QY 141 IleProAlaGluAspSerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160  
 Db 834 CTTCACAGGA---AGCAATTCCTTACGCCGATGCTCTATAT-----CCTATGATGAG 884  
 QY 161 GluLeuGlnGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLys---ThrGluGly 179  
 Db 885 ATGATCCAAAACCTTACGAAA-----TACGGGTTCAAGAAGACGAAAGAGA 929  
 QY 180 GlyValValAsnThr 184  
 Db 930 GGATGTTGTGAGACA 944  
 RESULT 8  
 AAF86431  
 ID AAF86431 standard: DNA; 349980 BP.  
 XX AAF86431;  
 AC 29-OCT-2001 (first entry)  
 XX  
 DE Pyrococcus abyssi genomic fragment #1.  
 XX  
 KM Hyperthermophilic archaeon; hyperthermophilic protein; ds.  
 XX  
 OS Pyrococcus abyssi.  
 XX  
 FH Key Location/Qualifiers  
 FT misc\_feature 300001..349980  
 FT /tag= a  
 FT /note= "This sequence overlaps with the 5' end of  
 AAH41223"  
 XX  
 XX FR2792651-Al.  
 XX  
 PD 27-OCT-2000.

PF 21-APR-1999; 99FR-00005034.  
 XX  
 PR 21-APR-1999; 99FR-00005034.  
 XX  
 XX (CNRS) CNRS CENT NAT RECH SCI.  
 PA (IFRE)- IFREMER INST FR RECH EXPL MER.  
 XX  
 XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O,  
 PI Quetellou J, Weissenbach J, Saurin W, Hellig R;  
 XX  
 XX WPI, 2001-126236/14.  
 DR  
 XX  
 PT New nucleotide sequences isolated from Pyrococcus abyssi encode proteins  
 PT useful in industry.  
 XX  
 PS Claim 1; Page 183-279; 1657pp; French.  
 XX  
 CC The present invention relates to the genomic sequence of Pyrococcus  
 CC abyssi and P. abyssi proteins (see AB96053-AB96642). P. abyssi is a  
 CC hyperthermophilic archaeon, which is isolated from deep-sea hydrothermal  
 CC vents. The present sequence is a fragment of the genomic sequence of P.  
 CC abyssi. The 3' end of this sequence overlaps with the 5' end of AAH4123.  
 CC The proteins of the present invention have various potential industrial  
 CC use, since the proteins are stable at very high temperatures, some up to  
 CC 110 degrees centigrade. Note: This patent is in the same parent family as  
 CC WO20005062, which contains additional sequences as shown in AB99132-  
 CC AAB99143, AAH75903-AAH75920 and AAG66436  
 CC  
 SQ Sequence 349980 BP; 95167 A; 78444 C; 78870 G; 97499 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 3.97e+03 Length: 349980  
 Score: 86.00 Matches: 40  
 Percent Similarity: 35.71% Conservative: 15  
 Best Local Similarity: 25.97% Mismatches: 41  
 Query Match: 8.68% Indels: 58  
 DB: 5 Gaps: 8  
 US-09-613-486-15 (1-198) x AAF86431 (1-349980)  
 QY 26 LysLeuLeuSerAlaGluValGluLysMetLeuValGluLysGlyAlaProAsnGluGly 45  
 Db 138056 AATTTGAGAGGCTTCATGATGTAATAAATGCTGAAGAAA---GTTCTGGGATAAC 138112  
 QY 46 IleGluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAlaArgThrTherProLys 65  
 Db 138113 ATTGACATATT----- 138124  
 QY 66 ValGlnArgAlaAspSerAspVal11ePheSerAsnSerPheGlyLysArgAsnVal--- 84  
 Db 138125 -----AAAGATGGGAGCGTGTCTATGCAAAAGGCTTGTGATACGAATATGTGAG 138175  
 QY 84 ----- 84  
 Db 138176 GCAAGGCTACCATCTACCCCGAGACATATACGAATTTGTTCAATAACCAAGAGCTTT 138235  
 QY 85 -----ValValThrGluGlyAspLeuLysValLeuAspGly 97  
 Db 138236 ACAGCTTGCAATATGAATTAAGTGAAGAGGCGGATTAAGT-----TTGATGAGC 138289  
 QY 98 CysAlaProLeuThrArgPheThrAsn---LysLeuArgThrPheGlyArgThrPheThr 116  
 Db 138290 -----CCGTTGAGAAATTCGTTAATATTAAGCTTACGCCCTTGGAGAACCAAGTGAC 138343  
 QY 117 GluAlaTyrValAspPheCysIleAlaTyrLysSHSLysLeuProGlnLeuAsnAlaAla 136  
 Db 138344 -----CTTCATACCTGTTACGCACTCTCGGGGATTCATCTCTAGATATAGCC 138394  
 QY 137 AlaGluLeu-----GlyIleProAlaGluAspSerTyrLeu 148  
 Db 138395 GAGGCTTCATAGACGAATGTCGCTGGGATACCTGTTG 138436  
 RESULT 9



CC tissue origin, which may be used in numerous diagnostic and prognostic  
 CC applications as well as cancer type-specific targets for therapeutic  
 CC intervention. The compounds that modulate the activity of a tumour  
 CC suppressor gene are useful in the treatment of cancer or as anti-cancer  
 CC drugs. The present sequence represents a polynucleotide of the invention.

XX Sequence 5311 BP; 1462 A; 1175 C; 1235 G; 1439 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	13.1	53.11	42	28	63	50
Percent Similarity:	85.00					
Best Local Similarity:	38.25%					
Query Match:	22.95%					
DB:	9					9

US-09-613-486-15 (1-198) x ADD29639 (1-5311)

QY 8 AsnLeuSerIleuValIleThrAspAlaSerSerIleuAsnGlyValAspIleuValLeu 27  
 DB 2030 AATATCCAAATATATCTT-----GCTCAAAACCTAAAGTGGGAGACGAAGAAGAAAG 2080  
 QY 28 LeuSerAlaGluVal-----GluIleMetIleuValGlnIleValAlaPheAsn 43  
 DB 2081 CTGACTTTGGAAATCGCCATTAGAGATTCACAGATTTCTCTACTGATGAGCCAC 2140  
 QY 44 GluGlyIleGluVal-----ValPheGlyLeuLeuLeuValAlaLeuAla 58  
 DB 2141 GCTGCTGCTGATCCCTTTTCAAGCACCGGCTGGATCTCTG----- 2185  
 QY 59 AlaArgThrThrSerProIleValGlnArgAlaAspSerAspValIlePheSerAsnSer 78  
 DB 2186 -----AAGAGCGCAGAGCAGACCCGTGCTGATCTCTG----- 2227  
 QY 79 Phe-----GlyGluArgAsnValValValThrGluGlyAspLeu 91  
 DB 2228 TTTCATGATGAGCGTGCATCCTGCTGATGAGAAAGTCTTCATCTCCATGAGAGCGCTG 2287  
 QY 92 LysIleValIleuAspGlyCysAlaPheLeuThrArgPheThrAsnIleuValArgThrPhe 111  
 DB 2288 AAG-----TGTCAGATCATCTCTCTTCTCGAAGAAAGAAATGGGAGCTT 2332  
 QY 112 GlyArgThrPheThrGluAlaIleValAspPheCysIle-----Ala 125  
 DB 2333 GGGATACCTTAAGTTGACAGCTGAAGAAGCTGTGTTCCAGAAACATCAGATGCTT 2392  
 QY 126 TyrIleHisIleuPheProGlnIleuAsnAlaAlaIleu-----Leu 139  
 DB 2393 GTTAAACGACATCCCTGACGAGCCAGACTATCAGCTTAAGGGGAGAAAGCTTCTTAT 2452  
 QY 140 GlyIleProAlaGluAspSerTyrIleuAlaAlaIleu-----LeuGlyThrCys 156  
 DB 2453 ACATTAACCTTAAGAAACACTTATAGATTCCAGAGCTTGGCAGAGCTTGCACAGCTGT 2512  
 QY 157 ProIleuVal 159  
 DB 2513 CCTGGACTG 2521

## RESULT 11

AAA91259  
 ID AAA91259 standard; DNA; 4766 BP.

AC AAA91259;

DT 06-AUG-2003 (revised)

DT 08-MAY-2001 (first entry)

DE GLRAV-5 nucleotide sequence.

KW GLRAV-5; grapevine leafroll virus; GLRAV infection; GLRAV coat protein;  
 KW GLRAV HSP70 homologue protein; viral gene mapping;  
 XX plant disease resistance; ds.

OS Grapevine leafroll virus.

XX Key Location/Qualifiers

FT Key 1..1593

FT CDS /tag= a

FT /product= "HSP70 homologue protein"

FT /partial /note= "No start codon given, Specifically claimed region"

FT CDS /tag= b

FT /product= "ORF 2 protein"

FT /note= "Specifically claimed region"

FT CDS /tag= c

FT /product= "coat protein"

FT /note= "Specifically claimed region"

FT CDS /tag= d

FT /product= "duplicate coat protein"

FT /note= "Specifically claimed region"

PN WO200105957-A2.

PD 25-JAN-2001.

PF 19-JUL-2000; 2000MO-US019708.

PR 19-JUL-1999; 99US-0144453P.

PA (AGRI-) AGRITOP INC.

PI Good XC, Montis J;

XX WPI, 2001-147339/15.

DR P-PSDB; AAY97688, AAY97689, AAY97690, AAY97691.

PT Novel grapevine leafroll virus polynucleotide useful as diagnostic and

PT probe, for viral gene mapping and for induced plant disease resistance.

PS Claim 1; Fig 1; 60pp; English.

XX This sequence represents a grapevine leafroll virus (GLRAV-5) DNA

CC sequence of the invention. The DNA sequence can be used in an expression

CC construct. The construct is useful for providing resistance to GLRAV

CC infection in a recombinant plant cell by transforming the plant cell with

CC it, where transcription of the polynucleotide sequence interferes with a

CC normal viral function such as movement, encapsidation or replication of

CC viral RNA. The polynucleotide sequence is expressed as an antisense

CC sequence and encodes a GLRAV coat protein, preferably a defective GLRAV

CC coat protein or a GLRAV HSP70 homologue protein. The GLRAV-5 DNA is

CC useful for the synthesis of GLRAV, as diagnostics and probes, for viral

CC gene mapping and for induced plant disease resistance. It is also useful

CC to detect and quantitate expression of GLRAV in plant tissue prior to use

CC in vegetative propagation, by detecting the presence of GLRAV RNA.

CC (Updated on 06-AUG-2003 to correct OS field.)

XX Sequence 4766 BP; 1329 A; 890 C; 1231 G; 1316 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	13	4766	32	40.46%	57	21
Percent Similarity:	84.50					
Best Local Similarity:	40.46%					
Query Match:	24.43%					
DB:	4					4

US-09-613-486-15 (1-198) x AAA91259 (1-4766)

QY 61 ThrThrSerProIleValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGly 80  
 DB 3663 TCACGCTCTCTTAAGGTGTCACGCTCTTCCATTCGACCTA--ACCGCTAAGTATGAT 3719



Db 817 ACTATGCGGAAAGTTGGGA 837

RESULT 13

ABA92787\_4/C

Continuation (5 of 7) of ABA92787 from base 400001 (Buchnera sp. genomic DNA SEQ ID NO:1)

WP Sequence split into 7 fragments LOCUS ABA92787 Accession ABA92787  
 WP Fragment Name Begin End  
 WP ABA92787\_1 1 110000  
 WP ABA92787\_2 110000 200001  
 WP ABA92787\_3 200001 310000  
 WP ABA92787\_4 310000 410000  
 WP ABA92787\_5 410000 510000  
 WP ABA92787\_6 510000 610000  
 WP ABA92787\_7 610000 640681

Alignment Scores:

Pred. No.: 1.57e+03 Length: 110000  
 Score: 83.50 Matches: 32  
 Percent Similarity: 37.18% Conservative: 26  
 Best Local Similarity: 20.51% Mismatches: 53  
 Query Match: 8.43% Indels: 45  
 DB: Gaps: 5

US-09-613-486-15 (1-198) x ABA92787\_4 (1-110000)

QY 65 LysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGlyGluArgAsnVal 84  
 Db 35915 AAAAATCAACCTGGATGATTAAGCATGATGACGAGGCTACTCCAGGTCAAAAGAAATAT 35855  
 QY 85 ValValThrGlnGlyAspLeuLysLysValLeuAsp----- 96  
 Db 35855 TTATTTGAAATGGAATATTAATAAATAATATGCAATAAATGCGCTTTAATG 35796  
 QY 97 -----GlyCysAlaPheLeuThrArg 103  
 Db 35795 GGTGTAACTGCTACCGGAAATGTCGTCGTGATCTTATTCGTCTAATGCTGCTCGT 35736  
 QY 104 PheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrValAspPheCys 123  
 Db 35735 ATGACTAAT-----ACCTAATG 35718  
 QY 124 IleAlaTyrLysHisLysLeuProGlnLeuAsnAlaIleAlaGluLeuGlyIleProAla 143  
 Db 35717 TTATCTGGAAGCTAATATGATGATATATTAAGCCGCTGACTATGGAATA----- 35664  
 QY 144 GluAspSerLysLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGln 163  
 Db 35663 -----TTCGCTGTAATTTTTCGGA-----GCACGAGTAGATATTACT 35625  
 QY 164 GlnSerArgLysMetPheAlaSerMetLysLysLysThrGlnGlyValValAsn 183  
 Db 35624 TCCGGAATAATTGATTTCTACTTCGGAAGCTATTATTAATAAATAATGAAAAATTGTT 35565  
 QY 184 ThrProValSerAsnLeuArgGlnLeuGlyArg---ArgGluValMet 198  
 Db 35564 ACACCAATTAATAAACACTACTCTCATAGATCAGGATTAAGAAATGATG 35517

RESULT 14

ACA47601 standard; DNA; 1773 BP.

XX ACA47601;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #29258.

XX Antisense; ds, prokaryotic essential gene; cell proliferation;

XX drug design; gene.

XX Staphylococcus haemolyticus.

PM WO200277183-A2.

XX 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-03429232.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GT, Yamamoto R, Forsyth RA, Xu HH;

PI P-PsDB; ABU43731.

DR WPI; 2003-029926/02.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 14; SEQ ID NO 35471; 1766bp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway (8) required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pat\_sequences

XX Sequence 1773 BP; 602 A; 305 C; 335 G; 531 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.64 Length: 1773  
 Score: 82.50 Matches: 35  
 Percent Similarity: 37.88% Conservative: 15  
 Best Local Similarity: 26.52% Mismatches: 43  
 Query Match: 8.32% Indels: 39  
 DB: Gaps: 6

US-09-613-486-15 (1-198) x ACA47601 (1-1773)

QY 4 MetSerAspSerAsnLeuSerAsnLeuValIleThrAspAlaSerSerLeu----- 20  
 Db 865 ATGATGCTCTCTATGCAAGATATGAGCTTAAACCGAATGATTTATTAATTA 924

```

QY 21 ---AsnGlyValAlaAspIysIysLeuSerAlaGluValGluLysMetLeuValGlnLys 39
DB 925 GGTAGTCGGTTTATGATTAATGACAACTG----- 954
QY 40 GlyAlaProAsnGluGlyIleGluValAlaPheGlyLeuLeuValTyrAlaLeuAla 59
DB 955 ---AATCCTATGAA-----TTTGCTCCTCATGCA 981
QY 60 ArgThrThrseryProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPhe 79
DB 982 ACCGTAGTCATGATGATATGATGATGAAATTAAACAGGTGATTCGACTGACTTA 1041
QY 80 GlyGluArgAsnValValAlaThrGluGlyAspLeuLysValLeuAspGlyCysAla 99
DB 1042 GGG-----ATAGTTCGCTGACTGTAATAAAGTACTCGAAGCTTATCT 1083
QY 100 ProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyr 119
DB 1084 -----CAAGATTACAAATAAA-----ACTTCCCATGAACTG 1119
QY 120 ValAspPheCysIleAlaTyrLysHisLysLeuPro 131
DB 1120 GTTAACTTGTATTGAAATTAAGCACAAACATCCT 1155

RESULT 15
AAA72313
ID AAA72313 standard; DNA; 1209 BP.
XX
XX AAA72313;
XX
XX 11-DEC-2000 (first entry)
XX
XX Erysipelothrix rhusiopathiae PPA fragment DNA.
XX
XX DE Protective polypeptide antigen; PPA; Fujisawa strain; swine erysipelas;
XX KM recombinant production; immunogenic; subunit vaccine;
XX KM transnucous administration; ds.
XX
XX OS Erysipelothrix rhusiopathiae.
XX
XX FH Location/Qualifiers
XX FT 1. 1209
XX FT CDS
XX FT /tag= a
XX FT /product= "Erysipelothrix rhusiopathiae 46.5 kD PPA
XX FT fragment"
XX FT /note= "No initiation or termination codon given in the
XX FT specification"
XX
XX JP3072345-B1.
XX
XX 31-UTL-2000.
XX
XX 31-MAR-1999; 99JP-00094004.
XX
XX 31-MAR-1999; 99JP-00094004.
XX
XX (NORO) NORINSUSANSHO KACHIKU EISEI SHIKENJOCHO.
XX PA (HGET) HIGETA SHOYU KK.
XX PA (FUOI-) FUJITA GAKUEN.
XX
XX WP1; 2000-551432/51.
XX DR P-PSDB; AAB14800.
XX
XX Novel recombinant protective polypeptide antigen useful as a vaccine for
XX PT protecting livestock against Erysipelothrix rhusiopathiae infection.
XX
XX Claim 3; Fig 4A-4C; 23pp; Japanese.
XX
XX This sequence represents DNA encoding a 46.5 kD immunogenic fragment of
XX CC the protective polypeptide antigen (PPA) of the bacterium Erysipelothrix
XX CC rhusiopathiae. This bacterium infects livestock, particularly pigs (swine
XX CC erysipelas), and is also able to cause disease in humans. The

```

```

CC recombinantly produced PPA fragment may be used as a subunit vaccine for
CC stimulating an immune response in animals against Erysipelothrix
CC rhusiopathiae. As the vaccine can be produced recombinantly, it can be
CC produced on a large scale. The vaccine is suitable for transnucous
CC administration
XX
SQ Sequence 1209 BP; 456 A; 177 C; 237 G; 339 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 4.36 Length: 1209
Score: 81.50 Matches: 33
Percent Similarity: 40.60% Conservative: 21
Best Local Similarity: 24.81% Mismatches: 50
Query Match: 8.22% Indels: 29
DB: Gaps: 5
US-09-613-486-15 (1-198) x AAA72313 (1-1209)
QY 60 ArgThrThrseryProLysValGlnArgAlaAspSerAspValIlePheSerAsnSerPhe 79
DB 358 AGATTAAAGATTCTCGAAGCTTGAAGACATACCTGATTTACTCGAATTACGAGGTG 417
QY 80 GlyGluArgAsnValValAlaThrGluGlyAspLeuLysValLeuAspGlyCysAla 99
DB 418 AAACACCGTATTATTAGTAAATATGAAAGTAAAGTTAA-----GGTAGAGCT 465
QY 100 ProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThr----- 116
DB 466 CCTTAGAAGCATTTATAGTTCCTCTTAAGACATGAAATTCGATGATGAAATTCCT 525
QY 117 -----GluAlaTyrValAspPheCysIleAla----- 125
DB 526 GCAGAAAGTAATTTATTCTGGAAGCCGATGAGATTTCTTAGTTTCAGATTCAAGCGAG 585
QY 126 TyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLysIleProAlaGluAsp 145
DB 586 TATTAATGACAACTAAATATATATCAACTTCTTGAGTCTAGGGGTGAGGAGTTTATT 645
QY 146 SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSer 165
DB 646 GACTAT-----AACCGGCTCGAAATATGATGAGAAAAA 678
QY 166 -----ArgLysMetPheAlaSerMetTyrAlaLeuLys 176
DB 679 GAACCTCATCCACTGATCTTGAATCTTATGCTATGCGG 717

```

Search completed: April 7, 2004, 02:20:39  
Job time : 394 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: April 7, 2004, 02:07:34 ; Search time 2353 Seconds  
(without alignments)  
2506.448 Million cell updates/sec

Title: US-09-613-486-15  
Perfect score: 991  
Sequence: 1 MEIMSDSNLSNLTVDASTL.....GGVNTPVSNLQGRREV 198

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Zgapop 6.0, Zgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+g2n.model -DEV=x1h  
-Q=/cgmt\_1/USPTO.spool/US09613486/runat\_05042004\_155121\_28274/app\_query.fasta\_1.391  
-DB=EST -QMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0  
-UNITS=bites -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09613486 @CGN 1.1 2135 @runat\_05042004\_155121\_28274 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEOUTRY -NES\_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOC  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THRAUS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hrc:\*  
9: gb\_estcl:\*  
10: gb\_estc2:\*  
11: gb\_hrc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_p1n:\*  
20: em\_gss\_vit:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mus:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vr1:\*  
28: qb\_gss1.\*

29: gb\_gss2.\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88.5	8.9	706	13 BQ986943	BQ986943 QGF10N21.
2	88.5	8.9	748	28 B2051926	B2051926 jms7h07.
3	87	8.8	477	12 BM132093	BM132093 TGEStx00
4	87	8.8	539	14 CP267035	CP267035 TGEStx02
5	85.5	8.6	562	13 BQ459495	BQ459495 H88017r
6	85.5	8.6	666	13 BU0007723	BU0007723 QGHE515.Y
7	85.5	8.6	671	12 BJ433495	BJ433495 B433495
8	85.5	8.6	704	28 B2434184	B2434184 BONFR91TF
9	85.5	8.6	713	9 AU034876	AU034876 AU034876
10	85.5	8.6	751	12 BU376355	BU376355 BU376355
11	85	8.6	654	13 BQ916284	BQ916284 OHR17G04.
12	84.5	8.5	610	13 BQ848646	BQ848646 QGA7M01.Y
13	84.5	8.5	611	13 BQ984758	BQ984758 QGESA05.Y
14	84.5	8.5	632	12 BG526441	BG526441 61-95 Ste
15	84.5	8.5	648	12 BG523153	BG523153 29-48 Ste
16	84.5	8.5	695	13 BQ993014	BQ993014 QGF27F14.
17	84	8.5	589	13 BQ983386	BQ983386 QGB16B23.
18	84	8.5	600	10 BF598403	BF598403 sv17b06.Y
19	84	8.5	732	28 BH976504	BH976504 odt67801.
20	84	8.5	1098	12 BG247808	BG247808 60235915
21	83.5	8.4	433	14 CA169078	CA169078 SCFBS101
22	83.5	8.4	633	28 BH096256	BH096256 RPCT-24-2
23	83.5	8.4	644	12 BU340622	BU340622 BU340622
24	83.5	8.4	644	12 BU343395	BU343395 BU343395
25	83.5	8.4	704	13 C92123	C92123 C92123 Dict
26	83.5	8.4	600	14 CB34251	CB34251 CA12EM000
27	83	8.4	755	29 CC722225	CC722225 OGVA288TH
28	83	8.4	911	12 BG421690	BG421690 602449789
29	82.5	8.3	661	28 AQ949879	AQ949879 Sheared D
30	82.5	8.3	762	14 CB569015	CB569015 AGENCOURT
31	82	8.3	607	12 BU775956	BU775956 f61h03.Y
32	82	8.3	627	13 BU026578	BU026578 OHG17R05
33	82	8.3	750	13 BQ481813	BQ481813 PV_GEB012
34	82	8.3	828	13 BU410316	BU410316 60158538
35	82	8.3	923	14 CA582924	CA582924 EST002599
36	82	8.3	1028	13 BU202935	BU202935 603102659
37	81.5	8.2	428	12 BG544734	BG544734 E2700 Ch1
38	81.5	8.2	560	28 CC155340	CC155340 CSU-K34.1
39	81.5	8.2	606	28 AZ079289	AZ079289 RPCT-23-4
40	81	8.2	513	28 AQ951071	AQ951071 Sheared D
41	81	8.2	727	28 BH929686	BH929686 odt98612.
42	81	8.2	807	28 B2577160	B2577160 msh2_5276
43	80.5	8.1	738	29 CG406315	CG406315 ZMWBB026
44	80.5	8.1	836	12 BU085103	BU085103 602870643
45	80	8.1	432	13 BU765335	BU765335 ssa816f01.

ALIGNMENTS

RESULT 1  
BQ986943  
LOCUS  
DEFINITION BQ986943 706 bp mRNA linear EST 21-AUG-2002  
ACCESSION QGF10N21, mRNA sequence.  
VERSION BQ986943  
KEYWORDS EST.  
SOURCE Lactuca sativa  
ORGANISM Lactuca sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; campanulids; Asterales; Ascleadeae; Cichoriaceae;

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Clitorieae; Lactuca.  
 1 (bases 1 to 706)  
 Kozik, R.A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L.,  
 Lin, H., van Damme, W., Lavelle, D., Chevalier, P., Ziegler, J.,  
 Ellison, P., Kolman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,  
 Lai, Z., Church, S., Jackson, J., and Bradford, K.  
 Lettuce and Sunflower ESTs from the Compositae Genome Project  
 http://compprojects.ucdavis.edu/  
 Unpublished (2002)  
 Contact: Alexander Kozik [R.W.Michelmore]  
 Department of Vegetable Crops, R.W.Michelmore Lab  
 University of California at Davis (UCD)  
 Asmundo Hall, UCD, Davis, CA 95616, USA  
 Tel: 1-(530)-742-1742  
 Fax: 1-(530)-752-9659  
 Email: akozik@ucdavis.org [michelmore@ucdavis.edu]  
 belongs to contig QC\_Ca\_contig3108, see http://cgdb.ucdavis.edu/  
 for details.  
 Plate: QGF10 row: N column: 21.

FEATURES  
 source

Location/Qualifiers  
 1..706  
 /organism="Lactuca sativa"  
 /mol\_type="mRNA"  
 /cultiVar="L.seriola"  
 /db\_xref="taxon:4236"  
 /clone="QGF10N21"  
 /lab\_host="E.coli"  
 /clone\_lib="QG\_EFGHJ lettuce seriola"  
 /note="Vector: pARCNOVA51AB. The library was constructed  
 from 10 different sources of RNA from a single genotype.  
 Separate cDNAs were generated using primers that  
 incorporated unique 5' and 3' tags to distinguish each  
 source of RNA. cDNAs were then pooled, size-fractionated,  
 directionally cloned into a custom medium-copy vector and  
 transformations made with four size classes to minimize  
 size bias. Details of each source of RNA and library  
 construction can be obtained at http://cgdb.ucdavis.edu/  
 TAG\_TISUS-germinating seeds  
 TAG\_TISUS-QG\_EFGHJ lettuce seriola  
 TAG\_SEQ=CTGTGCGG3"

## ORIGIN

## Alignment Scores:

Pred. No.: 1.45 Length: 706  
 Score: 88.50 Matches: 51  
 Percent Similarity: 41.36% Conservative: 28  
 Best Local Similarity: 26.70% Mismatches: 53  
 Query Match: 8.93% Indels: 19  
 DB: 13 Gaps: 13

US-09-613-486-15 (1-198) x BQ986943 (1-706)

QY 13 VallettraspalserseuenglyValaaplyslvleuLeuSerAlaGlVal 32  
 DB 137 GTTTGACCGAT---GGGCTTACCATGATGTTGCTCAGTTCCTGCTCAG--- 190  
 QY 33 GtlyuMeleuValGlnlyslValaProaenglylyleGlValaValaPhedlyLeu 52  
 DB 191 -----CTGCGACAGCAGAGGGGGTTTCGGAAGGTT----- 223  
 QY 53 leuLeuYrAlaLeuAlaAlaArghrThrSerProlys-----ValGlnArg 68  
 DB 224 -----AGGAAAGCTGCTCCAAAGATCTTGACCCGATTTGAGA 262  
 QY 69 AlaAsp-----SerAspValillePheSer----- 76  
 DB 263 GTTGAAGTTGTGACCGCTTGAAGAACATTGGGGGATGTGATTTGGGATTTGAAATCGAGA 322  
 QY 77 -----AsnSerPheGlyGlnArgAsnValValValThrGlnGlyAspLeuLys 92  
 DB 323 AGAGGGCAGATTAATATGCTTTGTGTGATTAAG-----CCTGGTGCATCT--- 364  
 QY 93 LysValLeuAspGlyCyAlaProLeuThr-----ArgPheThrAsnLysLeuArg 109

DB 365 AAGGTGTGATTTCTTGTCCTCCACTGCGGAATGTTGATGTAGACTTTAAG 424  
 QY 110 -----ThrpheGlyArg-----ThrpheThrGlnAlaYrValaAspPheCyAlaAlaYr 126  
 DB 425 GGAATGACAAAGAGAGAGCTTCTTACCAATGCAATTAATTAATGATGTGCTT 484  
 QY 127 LysHsLysLeuProGlnLeuAsnAlaAlaIaGlueuglyleProAlaGlnAspSer 146  
 DB 485 CAACATATTCAGAACAGCTTCTGCTGTAAGAAAGAGCTGCGACGTTGATCTTCA 544  
 QY 147 -----TyleuAla-----AlaAspPheLeuGlyThrCyAsp 157  
 DB 545 TCTTCATCTTGTGCTATTATGATGATGAGGGAATTTCTGATTTCACTTCAATGTCAA 604  
 QY 158 LysLeuSerGlnLeuGlnGlnSerArgLysMet 168  
 DB 605 TATAGCAAAACAAAAACAAACAAAGAT 637

RESULT 2  
 BZ051926/c 748 bp DNA linear GSS 09-OCT-2002  
 LOCUS jnr57h07.g1 B.oleracea001 Brassica oleracea genomic, genomic survey  
 DEFINITION  
 ACCESION BZ051926  
 VERSION BZ051926.1 GI:23652910  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Brassica oleracea  
 Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

1 (bases 1 to 748)  
 Deleantony, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,  
 Nash, W., Rabinowicz, P.D., and Wilson, R.K.  
 Whole genome shotgun reads from Brassica oleracea  
 Unpublished (2002)  
 Contact: Richard K. Wilson  
 Genome Sequencing Center  
 Washington University School of Medicine  
 Email: submissions@wustl.edu  
 Plate: jnr57 row: h column: 07  
 Seq primer: -28RPOT reverse  
 Class: shotgun  
 High quality sequence start: 56  
 High quality sequence stop: 551.  
 Location/Qualifiers

FEATURES  
 source

1..748  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:3712"  
 /clone\_lib="B.Oleracea001"  
 /note="Vector: pOTW13; whole genome shotgun library from  
 flowering buds. DNA was purified from a crude nuclear  
 prep using Brassica oleracea TO1000DH3 buds provided by  
 Thomas Osborn at the University of Wisconsin. Genomic  
 DNA was provided by Pablo Rabinowicz (CSH) and the  
 shotgun library prepared at Washington University Genome  
 Sequencing Center."

## ORIGIN

## Alignment Scores:

Pred. No.: 1.59 Length: 748  
 Score: 88.50 Matches: 33  
 Percent Similarity: 47.00% Conservative: 14  
 Best Local Similarity: 33.00% Mismatches: 40  
 Query Match: 8.93% Indels: 13  
 DB: 28 Gaps: 5

US-09-613-486-15 (1-198) x BZ051926 (1-748)

QY 84 ValValaValThrGlnGlyAspLeuLysLysValLeu-----AspGlyCyAlaProLeu 101  
 DB 84 ValValaValThrGlnGlyAspLeuLysLysValLeu-----AspGlyCyAlaProLeu 101

Db 551 ATACCATCGTGAAGACCACTACGACAAATCTTAGGAGTGTGAGTGCCTGACGATGCTC 492  
 Qy 102 ThArgPheThrAsnLysLeuAArgThrPheGlyAArgThrPheThrGluAlaValAsp 121  
 Db 491 ACTTTTAAGCTTAAACAAGACCACTATAGGCCCGATAGTGTCTGCAAGGTGACACAC 432  
 Qy 122 PheCysLeuAlaTyrLys--HisLysLeuProGlnLeuAsnAlaAlaGluLeuGly 140  
 Db 431 TTGATGCTCAAGGAGGAGCTCAAGATACAGGGTTCATGACGAGCTGCGGAGCT 372  
 Qy 141 ILleProAlaGluAspSerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160  
 Db 371 ATTAGTGCTATTGGAGTTTACTTACTCCCA-----GTAAGCACATCCCC--TTGACT 321  
 Qy 161 GluLeuGlnInserArgLysMetPheAlaSerMetTyrAlaLeuYsthrGlnGly 180  
 Db 320 CTGCAACACCAACCAAGAG-----CTAGAAGCTAAGGTGT 282

RESULT 3  
 BM132093 477 bp mRNA linear EST 27-NOV-2001  
 LOCUS T9ESTzyb0607.y1 T9RH Tachyzoite Subtracted cDNA Library Toxoplasma  
 DEFINITION gondii cDNA clone T9ESTzyb0607.y1 5', mRNA sequence.  
 VERSION BM132093  
 KEYWORDS EST.  
 SOURCE Toxoplasma gondii  
 ORGANISM Toxoplasma gondii  
 Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;  
 Sarcocystidae; Toxoplasma.  
 1 (bases 1 to 477)

## REFERENCE

1 (bases 1 to 477)  
 Tang, K., Cole, R., Fogarty, S., Sibley, L.D., Ajioke, J.A., White, M.,  
 Clifton, S., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, M.,  
 Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M.,  
 Ritter, E., Bennett, J., Franklin, C., Tsagaris, R., Ronko, I.,  
 Kennedy, S., Maguire, L., Waterston, R. and Wilson, R.  
 Toxoplasma EST Project  
 Unpublished (2001)  
 Contact: Clifton, S.  
 Toxoplasma EST Project  
 Washington University School of Medicine  
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: toxowatson.wustl.edu  
 Contact David Sibley (toxowest@borcim.wustl.edu) for further  
 information relating to organism, libraries, or clone availability.  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 420.  
 Location/Qualifiers

## TITLE

1. 477  
 /organism="Toxoplasma gondii"  
 /mol\_type="mRNA"  
 /strain="RH (Type 1)"  
 /db\_xref="taxon:5811"  
 /clone="T9ESTzyb0607.y1"  
 /dev\_stage="Tachyzoite"  
 /lab\_host="DH10B (GeneHog, Invitrogen, Inc.)"  
 /clone\_lib="T9RH Tachyzoite Subtracted cDNA Library"  
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:  
 XhoI; Toxoplasma RH strain tachyzoites were grown in human  
 foreskin fibroblast cultures in vitro. The library was  
 originally constructed by K.L.Wan, Cambridge University.  
 cDNAs were synthesized from polyA RNAs by oligo d(T)  
 priming and directionally cloned into the EcoRI to XhoI  
 sites of the lambda ZapII vector using the Zap-cDNA  
 synthesis kit (Stratagene). The primary cDNA library was  
 mass excised as phagemid using ExAssist helper phage  
 (Stratagene). Phagemid DNA was extracted by a pool of  
 phenol-chloroform method, and hybridized against a pool of  
 over-represented ESTs (N=12, from 5596 previous reads).  
 The subtracted library was electroporated into DH10B  
 (GeneHog, Invitrogen, Inc). WARNING: the library contains

## FEATURES

source

## ORIGIN

## Alignment Scores:

Pred. No.: 1.23 Length: 477  
 Score: 87.00 Matches: 30  
 Percent Similarity: 43.62% Conservative: 11  
 Best Local Similarity: 31.91% Mismatches: 33  
 Query Match: 8.78% Indels: 20  
 Db: 12 Gaps: 4

US-09-613-486-15 (1-198) x BM132093 (1-477)

Qy 88 GluGlyAspLeuLysLys--ValLeuAspGlyCysAlaProLeuThrArgPheThrAsn 106  
 Db 140 GAGGGGAGGTTGGGAGAGATAGTCTCTTGACAGTGTGCCCTTGAGGTGACACGAC 199  
 Qy 107 LysLeuArgThr-----PheGlyArgThr 114  
 Db 200 GAAAGCAGAGAGCGCGGATACACATTTCTCTTCAGTCTTCACTTTCAGGATG 259  
 Qy 115 PheThrGluAlaTyrValAspPheCysIle-----AlaTyrLysHis 128  
 Db 260 TTCTCCAGACACACCATGCTTATGATATCAAGTTGCTGGGGCTTTATGAAAG 319  
 Qy 129 LysLeuProGlnLeuAsnAlaAla--AlaGluLeuGlyIleProAlaGluAspSerTyr 147  
 Db 320 AAACGTCCCCGGTGCATGCTGCTTTTCGCAACAAATCTGAGCGTGAAGATCACAGA 379  
 Qy 148 LeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGlu 161  
 Db 380 CAGAGTACATGCACATCGCGCAATATGTCGGGACATCACTGA 421

RESULT 4  
 CF267035/c 539 bp mRNA linear EST 13-AUG-2003  
 LOCUS T9ESTzyj21c02.y1 T9MAS Tachyzoite cDNA Library Toxoplasma gondii  
 DEFINITION cDNA clone T9ESTzyj21c02.y1 5', mRNA sequence.  
 VERSION CF267035  
 KEYWORDS EST.  
 SOURCE Toxoplasma gondii  
 ORGANISM Toxoplasma gondii  
 Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;  
 Sarcocystidae; Toxoplasma.  
 1 (bases 1 to 539)

## REFERENCE

1 (bases 1 to 539)  
 Tang, K., Cole, R., Fogarty, S., Sibley, L.D., Ajioke, J.A., White, M.,  
 Clifton, S., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, K.,  
 Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M.,  
 Ritter, E., Bennett, J., Franklin, C., Tsagaris, R., Ronko, I.,  
 Kennedy, S., Maguire, L., Waterston, R. and Wilson, R.  
 Toxoplasma EST Project  
 Unpublished (2001)  
 Contact: Clifton, S.  
 Toxoplasma EST Project  
 Washington University School of Medicine  
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: toxowatson.wustl.edu  
 Contact David Sibley (toxowest@borcim.wustl.edu) for further  
 information relating to organism, libraries, or clone availability.  
 Seq primer: -40RP from Gibco.  
 Location/Qualifiers

## TITLE

1. 539  
 /organism="Toxoplasma gondii"  
 /mol\_type="mRNA"  
 /strain="Tachyzoite"  
 /db\_xref="taxon:5811"  
 /clone="T9ESTzyj21c02.y1"  
 /dev\_stage="Tachyzoite"

## FEATURES

source

1. 539  
 /organism="Toxoplasma gondii"  
 /mol\_type="mRNA"  
 /strain="Tachyzoite"  
 /db\_xref="taxon:5811"  
 /clone="T9ESTzyj21c02.y1"  
 /dev\_stage="Tachyzoite"

a small percentage of cDNAs derived from the human host  
 cells. Library Source: David Sibley, Washington  
 University."

```

/lab host="Electroten Blue cells (Stratagene)"
/clone.lib="mgms tachyzoite cdna library"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cdna library was constructed by Keliang Tang,
and Robert Cole at Washington University. cdna was
synthesized from Poly(A)+ mRNA using an oligo-d(T) primer
containing a XhoI site. Following second strand synthesis,
EcoRI adapters were ligated to the cdna, and products were
size-selected on sephacryl S500. The cdna were
directionally cloned into the EcoRI/XhoI prepared
pBluescript II SK+ vector, and electroporated into
Electroten Blue cells (Stratagene). The library may
contain a small percentage of host or bacterial
contaminants."

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 1.49 Length: 539
Score: 87.00 Matches: 30
Percent Similarity: 43.62% Conservative: 11
Best Local Similarity: 31.91% Mismatches: 33
Query Match: 8.78% Indels: 20
DB: 14 Gaps: 4

```

US-09-613-486-15 (1-198) x CF267035 (1-539)

```

QY 88 GAGGAGAGGTTGGGAGAGATAGTCTCTGACAGTGTCCCTTGAAGTACAGAAC 293
Db 352 GAGGAGAGGTTGGGAGAGATAGTCTCTGACAGTGTCCCTTGAAGTACAGAAC 293
QY 107 LysLeuArgThr-----PheGlyArgThr 114
Db 292 GAAACAGGAGAGTGGGAGATACACATTCTCCCTTCCAGTCTTCAGCTTTACAGATTG 223
QY 115 PheThrGluAlaValValAspPheCysIle-----AlaTyrLysHis 128
Db 232 TTCACAGCAGCACCATTGCTATTGTTATTCACAGTTCGCTGGGCTTTTGAAG 173
QY 129 LysLeuProGlnLeuAsnAlaAla---AlaGluLeuGlyLeuProAlaGluAspSerTyr 147
Db 172 AAACGTGCCCCGCTGCAATGCTGCTTTGCGCAACATACATGTAAGCTGAGATCAGAGA 113
QY 148 LeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGlu 161
Db 112 CAGAGTACATCGACATCGCGCAATGTCGGCGCATCGTGA 71

```

RESULT 5 BQ459495 562 bp mRNA linear EST 30-MAY-2002  
 LOCUS HA08017r HA Hordeum vulgare subsp. vulgare cdna clone HA08017  
 DEFINITION 5-PRIME, mRNA sequence.

ACCESSION BQ459495.1 GI:21267266

VERSION BQ459495

KEYWORDS EST

SOURCE Hordeum vulgare subsp. vulgare

ORGANISM Hordeum vulgare subsp. vulgare

REFERENCE Raddhak, V., Zhang, H., Weschke, W., Potokina, E. and Wobus, U.

AUTHORS Raddhak, V., Zhang, H., Weschke, W., Potokina, E. and Wobus, U.

TITLE Barley ESTs from developing seeds

JOURNAL Unpublished (2002)

COMMENT Molecular Markers Group, Department Genbank

Institute of Plant Genetics and Crop Plant Research (IPK)

Corrensstr. 3 06466, Gatersleben, Germany

Tel: 039462-5522

Fax: 039462-5595

Email: stein@ipk-gatersleben.de

Insert Length: 562 Std Error: 0.00

Plate: 8 row: 0 column: 17

Seq primer: M13rev.

## FEATURES

## source

## Location/Qualifiers

```

1..562
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="HA08017"
/tissue_type="embryosac"
/dev_stage="0-7 DAP (days after pollination)"
/lab_host="XU10-Gold"
/clone.lib="HA"
/note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of
cdna); Site 2: XhoI (3'-end of cdna); 0-7 DAP (days after
pollination) Due to a cloning artefact caused by the kit,
in most cases the EcoRI site is NOT present, as well as
the EcoRI adapter used for cloning. To excise the insert,
restriction sites upstream EcoRI should be used (e.g.
BamHI, SalI, PstI). NOTE: Also due to the cloning system
used Blue/white selection for recombinants is not 100%
reliable. Average insert size is 1 kb"

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 2.49 Length: 562
Score: 85.50 Matches: 40
Percent Similarity: 40.00% Conservative: 38
Best Local Similarity: 20.51% Mismatches: 84
Query Match: 8.63% Indels: 33
DB: 13 Gaps: 8

```

US-09-613-486-15 (1-198) x BQ459495 (1-562)

```

QY 22 GlyValAspLysLeuLeuSerAlaGluValGluLysMetLeuValGlnLysGlyAla 41
Db 4 GGCATGAGAGAGAGCCATGATGCTACACGAGGTCCT----- 45
QY 42 ProAsnGluGlyIleGluValValPheGlyLeuLeuLeuTyrAlaLeuAlaAlaArgThr 61
Db 46 ---GGAGAGGATCCAGAGTGAATTTGGGCTGTACTACATGATGCTGATCG 102
QY 62 ThrSerProLysValGlnArgAlaAspSerAspValIlePheSerAspPheGlyGlu 81
Db 103 CTGGAGAGAGAGCGGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 162
QY 82 -----ArgAsnValValValThrGluGlyAspLeuLysValLeuAspGly 97
Db 163 CATGACCTCCAGAGAGATGCTGTGAATCTGTGATTTTATGATGATGATGATGCT 222
QY 98 CysAlaProLeuThrArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGlu 117
Db 223 TACTGCGGTGTGAAGGATTCAGATGATGATGATGATGATGATGATGATGATGATGAT 282
QY 118 AlaTyrValAspPheCysIleAlaTyr-----LysHisLys 129
Db 283 AGGTGACACCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 342
QY 130 LeuProGlnLeuAsnAlaAlaGluLeu-----GlyTle--- 141
Db 343 CTT-----GTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 396
QY 142 ProAlaGluAspSerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSerGlu 161
Db 397 CCGTACGAAATACATTAATGCTTA-----CTCATGATGCTGCTTCAAGGTTATAT 450
QY 162 LeuGlnLysSerArgLysMetPheAlaSerMetTyr-----AlaLeuLysThrGluGly 179
Db 451 GTGATGAGAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 510
QY 180 GlyValValAsnThrProValSerAsnLeuArgGlnLeuGlyArg 194
Db 511 AATGCCCTTACAAAGATTAAGTGGCTGTGAGAGGTTGATAGG 555

```

RESULT 6  
 LOCUS BU007723  
 DEFINITION OGHSE15.ab1 OG FRGHJ lettuce serritola Lactuca sativa cDNA clone  
 OGHSE15, mRNA sequence.  
 ACCESSION BU007723  
 VERSION BU007723.1 GI:22442118  
 KEYWORDS EST.  
 ORGANISM Lactuca sativa  
 SOURCE Lactuca sativa  
 Lactuca sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; campanulids; Asterales; Asteraceae; Cichorioideae;  
 Cichorieae; Lactuca.  
 1 (bases 1 to 666)  
 Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Riesberg, L.,  
 Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, U.,  
 Ellison, P., Kolkmann, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,  
 Lai, Z., Church, S., Jackson, L. and Bradford, K.  
 Lettuce and Sunflower ESTs from the Compositae Genome Project  
 http://compositae.ucdavis.edu/  
 Unpublished (2002)  
 CONTACT: Alexander Kozik [R.W.Michelmore]  
 Department of Vegetable Crops, R.W.Michelmore lab  
 University of California at Davis (UCD)  
 Asmudson Hall, UCD, Davis, CA 95616, USA  
 Tel: 1-(530)-742-1742  
 Fax: 1-(530)-752-9659  
 Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]  
 belongs to contig OG\_CA.Contig3108, see http://cgpdb.ucdavis.edu/  
 for details.  
 Plate: OGH5 row: E column: 15.  
 Location/Qualifiers  
 1..666  
 /organism="Lactuca sativa"  
 /mol\_type="mRNA"  
 /cultivar="L.serritola"  
 /db\_xref="taxon:4236"  
 /clone="OGHSE15"  
 /lab\_host="E.coli"  
 /clone\_lib="OG\_FRGHJ lettuce serritola"  
 /note="Vector: pBRCDNAstf1B; The library was constructed  
 from 10 different sources of RNA from a single genotype.  
 Separate cDNAs were generated using primers that  
 incorporated unique 5' and 3' tags to distinguish each  
 source of RNA. cDNAs were then pooled, size-fractionated,  
 directionally cloned into a custom medium-copy vector and  
 transformations made with four size classes to minimize  
 size bias. Details of each source of RNA and library  
 construction can be obtained at http://cgpdb.ucdavis.edu/  
 TAG\_TISSUE=chemical induction  
 TAG\_LIB=OG FRGHJ lettuce serritola  
 TAG\_SEQ=TGTAGCCGGG"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3.25 Length: 666  
 Score: 85.50 Matches: 50  
 Percent Similarity: 40.78% Conservative: 23  
 Best Local Similarity: 27.93% Mismatches: 47  
 Query Match: 8.63% Indels: 59  
 DB: 13 Gaps: 13  
 US-09-613-486-15 (1-198) x BU007723 (1-666)  
 Oy 13 ValletThraepalaseSerLeuansglyValasplysleuLeuSerAlaGluVal 32  
 Db 184 GTTTGACCGAT---GGGTCTACCATGATGTGATCCAGATCTTGCGTTCCAG--- 237  
 Oy 33 GlulysMetLeuValGlnlyGlyAlaProAsnGluGlyIleGluValAlaPheGlyLeu 52  
 Db 238 -----CTGGCAGCACAGAGGGCGCTTTCGGAAGGTGT----- 270

Oy 53 LeuLeuTyzAlaLeuAlaAlaArgThrThrSerProlye-----ValGlnArg 68  
 Db 271 -----AGGAAGCTGCTCCAAAGATCTTGAGCGCATATGACG 309  
 Oy 69 AlaAsp-----SerAspValIlePheSer----- 76  
 Db 310 GTTGAAAGTTGTGACGCGCTGAGAACATTTGGGGAGATGATTTGGGATTTGAATTGACG 369  
 Oy 77 -----AsnSerPheGlyGluArgAsnValValValThrGluGlyAspLeuLys 92  
 Db 370 AGAGGCGAGATTATACCTTTGTTGATAG-----CCTGTGAGACT--- 411  
 Oy 93 LysValLeuAspGlyCysAlaProLeuThr-----ArgPheThrAsnLysLeuArg 109  
 Db 412 AAGGTGTGATTTCTTGTGCCACCTTCCGAAATGTTTCAGTATGTGATCTTAAAG 471  
 Oy 110 -----ThrPheGlyArg-----ThrPheThrGluAlaTyzValAspPheCysIleAlaTyz 126  
 Db 472 GGAATGCAAAAGGAGAGCTTCTTACACATGCAATTAATTAATTTGATTTGGCC 531  
 Oy 127 LysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAspSer 146  
 Db 532 CAACATTTGACAGACCGCTTCTGCTGTANAGAAAGACGTGTGGCAGCTTATCTTCA 591  
 Oy 147 -----TyrLeuAla-----AlaAspPheLeuGlyThrCys 156  
 Db 592 TCTTCATCTTGTGCTATTATGATAGAGGGAATTTTGTGATTTTCACTTCATATGT 648  
 RESULT 7  
 LOCUS BU433495/c 671 bp mRNA linear EST 13-MAR-2002  
 DEFINITION BU433495 Dictyostelium discoideum cDNA library, VF Dictyostelium  
 discoideum cDNA clone ddv22f05 3', mRNA sequence.  
 ACCESSION BU433495  
 VERSION BU433495.1 GI:19408217  
 KEYWORDS EST.  
 SOURCE Dictyostelium discoideum  
 ORGANISM Dictyostelium discoideum  
 Eukaryota; Mycetozoa; Dictyostelida; Dictyostellum.  
 REFERENCE 1 (bases 1 to 671)  
 Uruhashima, H., Tanaka, Y., Kohara, Y. and Shii, T.  
 Full length cDNA of Dictyostelium discoideum at the vegetative  
 stage  
 Unpublished (2002)  
 CONTACT: Tadasu Shin-i  
 Center for Genetic Resource Information  
 National Institute of Genetics  
 1111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tsbini@genes.nig.ac.jp.  
 Location/Qualifiers  
 1..671  
 /organism="Dictyostelium discoideum"  
 /mol\_type="mRNA"  
 /strain="AX4"  
 /db\_xref="taxon:44689"  
 /clone="ddv22f05"  
 /sex="mat A"  
 /dev\_stage="Growth phase"  
 /clone\_lib="Dictyostelium discoideum cDNA library, VF"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3.29 Length: 671  
 Score: 85.50 Matches: 46  
 Percent Similarity: 35.45% Conservative: 21  
 Best Local Similarity: 24.34% Mismatches: 85  
 Query Match: 8.63% Indels: 37  
 DB: 12 Gaps: 5  
 US-09-613-486-15 (1-198) x BU433495 (1-671)

	Dy	33	GATGTCAGTCTGAAGTTGGTGTAAAGACCAATTGAGAGATTGAGAGATTGGCGAGTATTA	604
	Dy	663	GAAAGAGTTAAGTTAGTTAGTGTAAAGACCAATTGAGAGATTGAGAGATTGGCGAGTATTA	604
	Qy	53	LeuLeuTyrrAlaLeuAlaAlaArgThrThrSerProLysValGlnArgAlaAspSerAsp	72
	Dy	603	TCATCATCATTGATTTGCAGCTCATATGACA-----CAATTACCAAGTGAAAGAT	556
	Qy	73	ValIlePheSerAsnSerPheGlyGluArgAsnValValValThrGluGlyAspLeuLys	92
	Dy	555	TTTGATTGTGACGAAAGAGTGGCTAATAAGTTGTTACATTGTCCAGAATCAAAATTTAAAG	496
	Qy	93	LysValLeuAspGlyCysAlaProLeuThrArgPhe-----	104
	Dy	495	TTGGGTGTTAAAGGATATTTCACCAGTTCATMAACTTTAAAACAAGGTGTAAACGTTTCA	436
	Qy	105	-----ThrAsnLysLeuArgThrPheGlyArgThrPheThr	116
	Dy	435	GTTTGCTACTGATTTGCCCGCTAGTAARGATGATTTAGTAGTGGTGGGAGATTACGTAAT	376
	Qy	117	GluAlaTyrrValAlaAspPheCysIleAlaTyrrLysHisLysLeuProGlnLeuAsnAlaAla	136
	Dy	375	GCACCTTACATTTGATAAACCTCTGCGAAATCACTCATTCATCAAT-----	334
	Qy	137	AlaGluLeuEngLyIleProAlaGluAspSerTyrrLeu-----	148
	Dy	333	---GAAAGGTGTGAACCTGTACACACATCATATAAGATTGTGTCAARGGCAACATATAAC	277
	Qy	149	AlaAlaAspPheLeuEngLyThrCysProLysLeuSerGluLeuGlnInserArgLysMet	168
	Dy	276	GGTCAAAAGGATTTGGGTATGATGATGATMAAGTTGATTCATCTCAAATTTGTAA-----	223
	Qy	169	PheAlaSerMetTyrrAlaLeuLysThrGluGlyValValAsnThrProValSerAsn	188
	Dy	222	TTTCAGATTTTCATGCGCTGCTCAAAAGTTTCAAGTCATCCAGTCTATGATCAATCTCTCAT	163
	Qy	189	LeuArgGlnLeuEngLyArgArgGluVal	197
	Dy	162	TTAGTTATGTTGGTACTAATCATCATGTC	136
RESULT 8	BZ434184/c			
JOCUS	DEFINITION	BZ434184	704 bp	DNA linear GSS 13-DEC-2002
ACCESSION	BOJFR91TF BO_1_6_2_KB_tot Brassica oleracea genomic clone BONFR91,			
VERSION	BZ434184			
KEYWORDS	BZ434184.1 GI:26684219			
SOURCE	GSS.			
ORGANISM	Brassica oleracea			
	Brassicaceae			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;			
	rosids; eurosids II; Brassicales; Brassicaceae; Brassica.			
REFERENCE	1 (bases 1 to 704)			
AUTHORS	Town,C.D., Van Aken,S., Utehrack,T., Koo,H. and Frazer,C.M.			
TITLE	Whole genome shotgun sequencing of Brassica oleracea			
JOURNAL	Unpublished (2001)			
COMMENT	Other_GSS: BONFR91TR			

## FEATURES

```
/organism="Brassica oleracea"  
/mol_type="genomic DNA"  
/strain="TO1000DH3"  
/db_xref="taxon:3712"
```

```

/clonea="BONFR91"
/clone1lb="BO 1.6-2 KB tot"
/note="Vector: pHS01.7 Site 1: BstXI; 1.6-2 Kb sheared
total DNA inserted into pHS01 using BstXI linkers"

```

Alignment Scores:	
Pred. No.:	3.55
Score:	85.50
Percent Similarity:	38.16%
Best Local Similarity:	23.67%
Query Match:	8.63%
DB:	28
US-09-613-486-15 (1-198) X BZA34484 (1-704)	
	Length: 704
	Matches: 49
	Conservative: 30
	Mismatches: 81
	Indels: 47
	Gaps: 8

Oy		7	SerAsnSerSerAsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAlaPheLys	26
Dd		574	TCTCACTTCAACAACAAAGTCATGCACAACTTGCCCTCATCATCTCAAGCCTTTGGAGACTCAA	515
Oy		27	Leu-----LeuSerAlaGluValGln	33
Dd		514	TTTGCTTCTAATGCCTTCATCTCCTCAAGCACGCCAATGGCGCTCTTACCAGGAATAATGAAA	455
Oy		34	LysMetLeuValGlnLysGlyAlaProAsnGlnGlyIleGluValAlaPheGlyLeuLeu	53
Dd		454	AAG-----AACCCCAAGAGTCCTGCAATGTGTCCTCTCC-----	419
Oy		54	LeuTrpAlaLeuAlaAlaArgThrThrSerProLysValGlnArgAlaAspSerAsp---	72
Dd		418	-----ACTACTCTCCAGAGATTGATGGATGCATCATGCACAA	380
Oy		73	-----ValIlePheSerAsnSerPheGlyGlu--ArgAsnVal	84
Dd		379	GAGAGAGATGAGATTGAAGACTGGTATTGGAACTGAGTTGGGGAGAAGTTGAGACTTT	320
Oy		85	ValValThrGlnGlyAspLeuLysLysValLeuAsnGlyLysAlaProLeuThrArgPhe	104
Dd		319	GTTGTGGCCACAGCTGTGAAGCACCAATTGTATGAGC---GCTCCAGAGAAGGTTGAAGCA	263
Oy		105	ThrAsnLysLeuArgThrPheGlyArgThrPheThrGlnAlaTryrValAspPheCysIle	124
Dd		262	ACAAATCTGCACAAAGCTGAGCACAA-----GCTGACAGAAACAAGTTGGAAAGAGACT	209
Oy		125	AlaTryrLysHisLysLeuProGlnLeuAsnAlaAlaIlaGluLeuGlyIleProAlaGlu	144
Dd		208	CACACACAGCTGAAAAAAGATTAACCTTAAAGAAAGCCATGAGGTGAGCATCAACCTAT	149
Oy		145	AspSerTyrlleuAlaAlaAspPheLeuGlyThyrCysProLysLeuSerGlnLeuGlnGln	164
Dd		148	GATAGACTCTCTTTCCCCCAAGAGTCTCACCAAGCTCAAGACGAAGTGCTCTCCAAG	89
Oy		165	SerArgLysMetPheAlaSerMetTyrlaLeuLysThrGlnGlyGlyValAlaAsnThr	184
Dd		88	TTCAAGAAAGATCTTAGATGATT-----GGGTCACAGGCTTCCA	50
Oy		185	ProValSerAsnLeuArgIn	191
Dd		49	GAATCTCGAGTATGCGGTGA	29

RESULT 9	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
AU034876			AU034876						
		713 bp							
		(H.Urushihara)							
		linear							
		Dictyostelium							
		discoidium							
		CDNA clone SL8629, mRNA sequence.							
			AU034876						
			AU034876.1	GI:3800300					
			EST.						
			Dictyostelium discoidium						
			Dictyostelium discoidium						
			Dictyostelium discoidium						
			Eukaryota; Mycetozoa; Dictyostelid; Dictyostelium.						
			1 (bases 1 to 713)						
			Morita,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,						

TITLE Yoshino, R., Mittra, B.N., Pi, M., Sato, T., Takemoto, K., Yasukawa, H., Williams, J., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.  
The Dictyostelium developmental cDNA project: generation and analysis of expressed sequence tags from the first-finger stage of development

JOURNAL DNA Res. 5 (6), 335-340 (1998)

MEDLINE 99156227

PubMed 10048482

## COMMENT

Contact: Hideko Urushihara  
Institute of Biological Sciences  
University of Tsukuba  
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan  
Tel: 81-298-53-4664  
Fax: 81-298-53-6614  
Email: hideko@biol.tsukuba.ac.jp  
PROJECT = 'Dictyostelium discoideum cDNA project in Japan'  
POLYA-No.

## FEATURES

source

Location/Qualifiers  
1..713  
/organism="Dictyostelium discoideum"  
/mol\_type="mRNA"  
/strain="AX4"  
/db\_xref="taxon:44689"  
/clone="SLE629"  
/dev\_stage="slug"  
/clone\_lib="Dictyostelium discoideum SL (H. Urushihara)"

## ORIGIN

Alignment Scores:

Pred. No.: 3.62 Length: 713  
Score: 85.50 Matches: 46  
Percent Similarity: 35.45% Conservative: 21  
Best Local Similarity: 24.34% Mismatches: 85  
Query Match: 8.63% Indels: 37  
DB: Gaps: 5

US-09-613-486-15 (1-198) x AU034876 (1-713)

QY 33 GltuysMeLeuValGlnLysGlyAlaProAsnGluGlyIleGluValValPheGlyLeu 52  
Db 40 GAAGAGTTAAGGTTAGTGTAGAGACCAATTGAGAGATTGAGAGATTGGAGTATTA 99  
QY 53 LeuLeuTyraLeuAlaAlaArgThrThrserrProLysValGlnArgAlaAspSerAsp 72  
Db 100 TCATCATTCATTGATTGACGTGCATATGACA-----CAATTAAACAAGTGAAGAT 147  
QY 73 ValIlePheSerAsnSerPheGlyGluArgAsnValValValThrGluGlyAspLeuLys 92  
Db 148 TTAGATTGACGACCAAGAGTGTATTATTTAGTTGACTTGTCCAGATCAAAATTAAAG 207  
QY 93 LysValLeuAspGlyCysAlaProLeuThrArgPhe----- 104  
Db 208 TTGGGTGTTAAAGGATTTTACACAGTTCACTTAACCTTTAAACAAGGTAAAGTTTCA 267  
QY 105 -----ThrAsnLysLeuArgThrPheGlyArgThrPheThr 116  
Db 268 GTTGCTACTGATTCGCCCGCTAGTAAATATGATTAGTATGTTGGGTGAATTACGTA 327  
QY 117 GluAlaTyraValAspPheCysIleAlaTyrlsHisLysLeuProGlnLeuAsnAlaAla 136  
Db 328 GCAGCTTACATTGATTAACCTCTGCAAAATACCTATTCATTT----- 369  
QY 137 AlaGluLeuGlyIleProAlaGluAspSerTyLeu----- 148  
Db 370 ---GAAGGTGGGAACCTGTAACACCATCATTAAGATTGTCATAGGCACTATTAAAC 426  
QY 149 AlaAlaAspPheLeuGlyThrCysProLysLeuSerGluLeuGlnGlnSerArgLysMet 168  
Db 427 GGTGGAAGGAGCATTTGGTATCGATGATAAGTTGGTTCATTCAAAATTGGTAA----- 480  
QY 169 PheAlaSerMetTyraLeuLysThrGluGlyGlyValValAsnThrProValSerAsn 188  
Db 481 TTGCAAGATTTCATCGCTGTCAAAAGTTTCAAGTCACTCAAGTCTATGATCAATCTTCAT 540

QY 169 LeuArgGlnLeuGlyArgArgGluVal 197  
Db 541 TTAGTTAAGTTGGTACTAATCATGTC 567

RESULT 10

BU376355/c

LOCUS BU376355

DEFINITION BU376355 Dictyostelium discoideum cDNA library, CF Dictyostelium

ACCESSION BU376355

VERSION BU376355.1 GI:19285738

KEYWORDS EST.

SOURCE Dictyostelium discoideum

ORGANISM Dictyostelium discoideum

REFERENCE Eukaryota; Mycetozoa; Dictyostellidae; Dictyostelium.

AUTHORS 1 (bases 1 to 751)

TITLE Urushihara, H., Tanaka, Y., Kohara, Y. and Shin-i, T.

Full length cDNA of Dictyostelium discoideum at the culmination

stage

JOURNAL Unpublished (2002)

COMMENT Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1..751  
/organism="Dictyostelium discoideum"  
/mol\_type="mRNA"  
/strain="AX4"  
/db\_xref="taxon:44689"  
/clone="ddc28n02"  
/sex="mat A"  
/dev\_stage="Quinmation stage"  
/clone\_lib="Dictyostelium discoideum cDNA library, CF"

## ORIGIN

Alignment Scores:

Pred. No.: 3.92 Length: 751  
Score: 85.50 Matches: 46  
Percent Similarity: 35.45% Conservative: 21  
Best Local Similarity: 24.34% Mismatches: 85  
Query Match: 8.63% Indels: 37  
DB: Gaps: 5

US-09-613-486-15 (1-198) x BU376355 (1-751)

QY 33 GltuysMeLeuValGlnLysGlyAlaProAsnGluGlyIleGluValValPheGlyLeu 52  
Db 663 GAAGAGTTAAGGTTAGTGTAGAGACCAATTGAGAGATTGAGAGATTGGAGTATTA 604  
QY 53 LeuLeuTyraLeuAlaAlaArgThrThrserrProLysValGlnArgAlaAspSerAsp 72  
Db 603 TCATCATTCATTGATTGACACTCATATGACA-----CAATTAAACAAGTGAAGAT 556  
QY 73 ValIlePheSerAsnSerPheGlyGluArgAsnValValValThrGluGlyAspLeuLys 92  
Db 555 TTAGATTGACGACCAAGAGTGTATTATTTAGTTGACTTGTCCAGATCAAAATTAAAG 496  
QY 93 LysValLeuAspGlyCysAlaProLeuThrArgPhe----- 104  
Db 495 TTGGGTGTTAAAGGATTTTACACAGTTTAATAAAGTTTAAACAAGGTAAAGTTTCA 436  
QY 105 -----ThrAsnLysLeuArgThrPheGlyArgThrPheThr 116  
Db 435 GTTGCTACTGATTCGCCCGCTAGTAAATATGATTAGTATGTTGGGTGAATTACGTA 376  
QY 117 GluAlaTyraValAspPheCysIleAlaTyrlsHisLysLeuProGlnLeuAsnAlaAla 136  
Db 375 GCAGCTTACATTGATTAACCTCTGCAAAATACCTATTCATTT----- 334

QY 137 AlGluLeuGlyIleProAlaGluAspSerTyrLeu----- 148  
 Db 333 ---GAAGGTGGTAACCTGTACACCATCATTAAGATTGTCGAATGGCACTATTAC 277  
 QY 149 AlaAlaaspheLeuGlyThrCysProlyLeuSerGluLeuGlnInserArgLysMet 168  
 Db 276 GGTGCAAGGCGATTGGTATCGATGATGAATAGCTGTTCACCTTCAATGGTAA----- 223  
 QY 169 PheAlaSerMetTyrAlaLeuGlyThrGluGlyValAlaSerThrProValSerAsn 188  
 Db 222 TTTCGACATTTCATCGCTGTCGCAAGTTTCAAGTCATCCAGTCATGATCCAACTCTCAT 163  
 QY 189 LeuArgGlnLeuGlyArgArgGluVal 197  
 Db 162 TTAGTTATGTGTGTACTATCATGTC 136  
 RESULT 11  
 BQ916284 654 bp mRNA linear EST 19-AUG-2002  
 LOCUS OHB17G04.YG.ab1 QG ABCDI sunflower RHA801 Helianthus annuus cDNA  
 DEFINITION clone OHB17G04, mRNA sequence.  
 ACCESSION BQ916284  
 VERSION BQ916284.1 GI:22315065  
 SOURCE EST.  
 ORGANISM Helianthus annuus (common sunflower)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asterids; campanulids; Asterales; Asteraceae; Asteroideae;  
 Heliantheae; Helianthus.  
 1 (bases 1 to 654)  
 Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,  
 Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,  
 Ellison,P., Kolkmann,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,  
 Lai,Z., Church,S., Jackson,L. and Bradford,K.  
 Lettuce and Sunflower ESTs from the Compositae Genome Project  
 http://compgenomics.ucdavis.edu/  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Alexander Kozik [R.W.Michelmore]  
 Department of Vegetable Crops, R.W.Michelmore Lab  
 University of California at Davis (UCD)  
 Amundson Hall, UCD, Davis, CA 95616, USA  
 Tel: 1-(530)-742-1742  
 Fax: 1-(530)-752-9659  
 Email: akozik@ucdavis.org [michelmore@vegmail.ucdavis.edu]  
 belongs to contig QH\_CA\_Config3046, see http://cgpdb.ucdavis.edu/  
 for details.  
 Plate: QHB17 row: G column: 04.  
 FEATURES  
 source 1..654  
 location/Qualifiers  
 /organism="Helianthus annuus"  
 /mol\_type="mRNA"  
 /cultiivar="RHA801"  
 /db\_xref="taxon:4232"  
 /clone="OHB17G04"  
 /lab\_host="E.coli"  
 /note="Vector: pBRCNDS1AB: The library was constructed  
 from 11 different sources of RNA from a single genotype.  
 Separate cDNAs were generated using primers that  
 incorporated unique 5' and 3' tags to distinguish each  
 source of RNA. cDNAs were then pooled, size-fractionated,  
 directionally cloned into a custom medium-copy vector and  
 transformations made with four size classes to minimize  
 size bias. Details of each source of RNA and library  
 construction can be obtained at http://cgpdb.ucdavis.edu/  
 TAG\_TIS05=shoots environmental stress  
 TAG\_LIB=QH ABCDI sunflower RHA801  
 TAG\_SEQ=TCGACACGGG"

Pred. No.: 3.67 Length: 654  
 Score: 85.00 Matches: 41  
 Percent Similarity: 39.78% Conservative: 31  
 Best Local Similarity: 22.65% Mismatches: 89  
 Query Match: 8.58% Indels: 20  
 DB: 13 Gaps: 6

US-09-613-486-15 (1-198) x BQ916284 (1-654)

QY 11 AsnLeuValIleThrAspAlaSerSerLeuAsnGlyValAspLysLysLeuSerAla 30  
 Db 23 GATCTTGAGTCGCGAGAGTAAGCCGTTATTTCGGTTATATGCAAA---GGCTCGCT 79  
 QY 31 GluValGluLysMetLeuValGlnLysGlyAlaProAsnGlnGlyIleGluValAlaPhe 50  
 Db 80 TCTGTAAAGCTATGCTGACATATAGGCGCTGTAATAACGAACTATTAAGTTATATAC 139  
 QY 51 GlyLeuLeuLeuTyrAlaLeuAlaAlaArg-----ThrThrSer 63  
 Db 140 GATCTTCTTGATGATGCCGAGCGCAATGAGGACTATGACCCGTTGAGATCAA 199  
 QY 64 ProlysValGlnArgAlaAspSerAspValIlePheSerAsnSerPheGlyGlu---Arg 82  
 Db 200 GTGCCGATTGGTTCGCGACAGCTTCGGCGGTGTTAGCACGCGCAGTGGCGATTGCC 259  
 QY 83 AsnValValAlaThrGluGlyAspLeuLysValLeuAspGlyCysAlaProLeuThr 102  
 Db 260 GGATCGATGTGATCTGAAGGA-----AAAGTACAAAAGATTGGCGATT----- 304  
 QY 103 ArgPheThrAsnLysLeuArgThrPheGlyArgThrPheThrGluAlaTyrValAspPhe 122  
 Db 305 -----AGGTTCAAGCAAGAAAGAAACACTTTTGTGTGTGTTCTGATCT 352  
 QY 123 CysIleAlaTyrLysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIlePro 142  
 Db 353 TTGAGACGATCAAGAAAGATTTAAGAGGTAACGGCGAGCTAGATGTGTTCCGG 412  
 QY 143 AlaGluAsp-----SerTyrLeuAlaAlaAspPheLeuGlyThrCysProLysLeuSer 160  
 Db 413 GTTGATGATTTATATGAAATGGAAGAGGTATGTATCGAAAGCTTTAATTCATTCCAG 472  
 QY 161 GluLeuGlnGlnSerArgLysMetPheAlaSerMetTyrAlaLeuLysThrGluGly 180  
 Db 473 AAAAGCGAACAACCTGAAAGACATCGATCATGACAGCAGCATTTAAAGAACTGGA 532  
 QY 181 Val 181  
 Db 533 ATC 535  
 RESULT 12  
 BQ848646 610 bp mRNA linear EST 14-AUG-2002  
 LOCUS OGATW01.YG.ab1 QG ABCDI lettuce salina lactuca cDNA clone  
 DEFINITION OGATW01, mRNA sequence.  
 ACCESSION BQ848646  
 VERSION BQ848646.1 GI:22234115  
 SOURCE EST.  
 ORGANISM Lactuca sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; campanulids; Asterales; Asteraceae; Cichorioideae;  
 Cichorieae; Lactuca.  
 1 (bases 1 to 610)  
 Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,  
 Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,  
 Ellison,P., Kolkmann,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,  
 Lai,Z., Church,S., Jackson,L. and Bradford,K.  
 Lettuce and Sunflower ESTs from the Compositae Genome Project  
 http://compgenomics.ucdavis.edu/  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Alexander Kozik [R.W.Michelmore]  
 Department of Vegetable Crops, R.W.Michelmore Lab



University of California at Davis (UCD)  
 Asmudson Hall, UCD, Davis, CA 95616, USA  
 Tel: 1-(530)-742-1742  
 Fax: 1-(530)-752-9659  
 Email: akozik@ucdavis.org [michelmore@vegmail.ucdavis.edu]  
 belongs to contig QG\_Ca\_Contig3108, see http://cspdb.ucdavis.edu/  
 for details  
 Plate: QG47 row: M column: 01.

## FEATURES

source

Location/Qualifiers  
 1. 610  
 /organism="Lactuca sativa"  
 /mol\_type="mRNA"  
 /cultivar="Salinas"  
 /db\_xref="taxon:4236"  
 /clone="QG47M01"  
 /lab\_host="E.coli"  
 /clone\_lib="QG\_ABCDI lettuce salinas"  
 /note="Vector: pBRCDNA5flab; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cspdb.ucdavis.edu/  
 TAG\_TISSUE=shoots environmental stress  
 TAG\_LIB=QG\_ABCDI lettuce salinas  
 TAG\_SEQ=TCGCACGGG"

## ORIGIN

## Alignment Scores:

Pred. No.: 3.82 Length: 610  
 Score: 84.50 Matches: 50  
 Percent Similarity: 40.78% Conservative: 23  
 Best Local Similarity: 27.93% Mismatches: 47  
 Query Match: 8.53% Indels: 59  
 DB: 13 Gaps: 13

US-09-613-486-15 (1-198) x BQ848646 (1-610)

QY 13 ValIleThrsAlaSerSerLeuAsnGlyValAspLysLeuLeuSerAlaGluVal 32  
 Db 137 GTTTGACCGAT---GGGTCTTACCATGATGATGATCCAGTCTCTGCGTCCAG---190  
 QY 33 GluLysMetLeuValGlnGlyAlaProAsnGluGlyTLeuValValPheGlyLeu 52  
 Db 191 -----CTGCACGACGAGGGCGTTTCGGGAAGTGT-----223  
 QY 53 LeuLeuTyzAlaLeuAlaAlaArgThrSerProLys-----ValGlnArg 68  
 Db 224 -----AGGAAAGCTGCTCCAAAGTCTTGGAGCCGATTAGAGG 262  
 QY 69 AlaAsp-----SerAspValIlePheSer-----76  
 Db 263 GTTGAATTGTGACGCTGAGAAACATTTGGGGGATGTGATGGGATTTGAATTCGGGA 322  
 QY 77 -----AnSerPheGlyGluArgAsnValValThrGluGlyAspLeuLys 92  
 Db 323 AGAGGGCAGATTATATACCTTTGGATAG-----CCTGGTGCATTT---364  
 QY 93 LysValLeuAspGlyCysAlaProLeuThr-----ArgPheThrAsnLysLeuArg 109  
 Db 365 AAGGTGTGATTTTGTTCCTCCCTTCGGAATGTTTCAAGATGAGTACTTTAAG 424  
 QY 110 -----ThrPheGlyArg-----ThrPheThrGluAlaArgValAspPheCysIleAlaArg 126  
 Db 425 GGAATGCAAAAGAGAGAGCTTTTACACATGCAATGCAATTAATTGATGTGCTT 484  
 QY 127 LysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyLeuProAlaGluAspSer 146  
 Db 485 CACATATTCAGAACAGCTTTCTGCTGCTCCAAAGAGAGAGCTGTGAGAGCTTATCTTCA 544

QY 147 -----TyrLeuAla-----AlaAspPheLeuGlyThrCys 156  
 Db 545 TCTTCATCTTTGTGCTCTTTTATGATGATAGAGGAGATTTTCTGATTTTCACTTCATATGT 601

## RESULT 13

BQ984758

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

ORGANISM

SOURCE

Lactuca sativa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Chloroidae; Chloridae; Lactuca.

1 (bases 1 to 611)

Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkman, J., Stabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.

Lettuce and Sunflower ESTs from the Composite Genome Project

http://compenomics.ucdavis.edu/

Unpublished (2002)

Contact: Alexander Kozik [R.W.Michelmore]

Department of Vegetable Crops, R.W.Michelmore Lab

University of California at Davis (UCD)

Asmudson Hall, UCD, Davis, CA 95616, USA

Tel: 1-(530)-742-1742

Fax: 1-(530)-752-9659

Email: akozik@ucdavis.org [michelmore@vegmail.ucdavis.edu]

belongs to contig QG\_Ca\_Contig3108, see http://cspdb.ucdavis.edu/  
 for details.  
 Plate: QG85 row: A column: 05.

## FEATURES

source

Location/Qualifiers  
 1. 611  
 /organism="Lactuca sativa"  
 /mol\_type="mRNA"  
 /cultivar="L. serriola"  
 /db\_xref="taxon:4236"  
 /clone="QG85A05"  
 /lab\_host="E.coli"  
 /clone\_lib="QG\_EFGHU lettuce serriola"  
 /note="Vector: pBRCDNA5flab; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cspdb.ucdavis.edu/  
 TAG\_TISSUE=chemical induction  
 TAG\_LIB=QG\_EFGHU lettuce serriola  
 TAG\_SEQ=TCGACCGGG"

## ORIGIN

## Alignment Scores:

Pred. No.: 3.83 Length: 611  
 Score: 84.50 Matches: 50  
 Percent Similarity: 40.78% Conservative: 23  
 Best Local Similarity: 27.93% Mismatches: 47  
 Query Match: 8.53% Indels: 59  
 DB: 13 Gaps: 13

US-09-613-486-15 (1-198) x BQ984758 (1-611)  
 QY 13 ValIleThrsAlaSerSerLeuAsnGlyValAspLysLeuLeuSerAlaGluVal 32  
 Db 137 GTTTGACCGAT---GGGTCTTACCATGATGATGATCCAGTCTCTGCGTCCAG---190

QY 33 GlnUleuMetLeuValGlnUlysglyAlaProAsnGlyIleGluValValPheGlyLeu 52  
 DB 191 -----CTGGCAGACGACGAGCGCTTCCGGAGAGTCTT-----ValGlnArg 68  
 QY 53 LeuLeuTyralaLeuAlaIaArgThrThrserProlys-----ValGlnArg 68  
 DB 224 -----AGGAAAGCTGCTCCAAAGATCTTGAGAGCCGATTAAGATA 262  
 QY 69 AlaAsp-----SerAspValIlePheSer-----76  
 DB 263 GTTGAAGTGTGACGCTGGAAGAACATTTGGGGATGTGATGGGATTTGAATTCAGAGA 322  
 QY 77 -----AnserPheGlyGluArgAsnValValValThrGluGlyAspLeuLys 92  
 DB 323 AGAGGCGAGATTATATAGCTTTGGATATAC-----CCTGGTGACCTT---364  
 QY 93 LysValLeuAspGlyCysAlaProLeuThr-----ArgPheThrAsnLysLeuArg 109  
 DB 365 AAGGTGGATCTTCTTGGTCCACTGCGGAAATGTTCAAGATGAGTACTTTAAG 424  
 QY 110 -----ThrPheGlyArg-----ThrPheThrGluAlaIaTyralaAspPheCysIleAlaTy 126  
 DB 425 GGAATGCAAAAGGAGAGCTCTTACACATGCAATGCAATTAATTTGATGTGGCT 484  
 QY 127 LysHisLysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAlaGluAspSer 146  
 DB 485 CAACATATTCAGAACCGACTTCTGCTGCTAAAGAAAGACGTGGCAGCTTATCTCA 544  
 QY 147 -----TyrlaAla-----AlaAspPheLeuGlyThrCys 156  
 DB 545 TCTTCATCTTGTCTCTTATTAATGATAGGGAATTTCTGATTTTCACTTCATATGT 601

RESULT 14  
 LOCUS BG526441 632 bp mRNA linear EST 16-NOV-2001  
 DEFINITION 61-95 Stevia field grown leaf cDNA Stevia rebaudiana cDNA 5', mRNA  
 sequence.  
 ACCESSION BG526441  
 VERSION BG526441.1 GI:16949914  
 KEYWORDS EST.  
 SOURCE  
 ORGANISM Stevia rebaudiana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; campanulids; Asterales; Asteraceae; Asteroidae;  
 Eupatoriaceae; Stevia.  
 1 (bases 1 to 632)  
 Brande, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.  
 Leaf ESTs from Stevia rebaudiana: A resource for gene discovery in  
 diterpene synthesis  
 Unpublished (2001)  
 JOURNAL Contact: Jim Brande  
 Genomics and Biotechnology  
 Agriculture and Agri-Food Canada - SCPPRC  
 1391 Sandford St., London, Ontario, CANADA, N5V 4T3  
 Tel: 519 457 1470  
 Fax: 519 457 3997  
 Email: brandeje@em.agr.ca  
 Seg primer: T3 promoter primer.  
 FEATURES  
 source  
 1..632  
 /organism="Stevia rebaudiana"  
 /mol\_type="mRNA"  
 /strain="751/1501"  
 /cultivar="Landrace"  
 /db\_xref="taxon:55670"  
 /tissue\_type="leaf"  
 /dev\_stage="field grown, mid-size"  
 /lab\_host="E. coli strain XL0LR"  
 /clone\_id="Stevia field grown leaf cDNA"  
 /note="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI; This  
 cDNA library was constructed from polyA+ enriched mRNA  
 from field grown leaves. Mid-size actively growing leaves

ORIGIN  
 Alignment Scores:  
 Pred. No.: 4.04 Length: 632  
 Score: 84.50 Matches: 48  
 Percent Similarity: 40.57% Conservative: 23  
 Best Local Similarity: 27.43% Mismatches: 47  
 Query Match: 8.53% Indels: 57  
 DB: 12 Gaps: 12

US-09-613-486-15 (1-198) x BG526441 (1-632)

QY 2 GlnUleuMetSerAspSerAsnLeuSerAsnLeu-----ValIle 14  
 DB 62 GAGTCAATGTCTTAATGAGATCTGCGGGCTTCCCTGTGTGACCTTCGTGCGGTGTG 121  
 QY 15 ThrAspAlaSerSerLeuAsnGlyValaAspLysLysLeuSerAlaGluValGluLys 34  
 DB 122 ACCGAT---GGGTCTATACCATGATGTTACTCGACTGCTACCTCGCTCCAG---169  
 QY 35 MetLeuValGlnUlysglyAlaProAsnGlyIleGluValValPheGlyLeuLeu 54  
 DB 170 ---CTGGCGGGCTCGAGAGCTTCCGTGAAGAGAT---202  
 QY 55 TyralaLeuAlaIaArgThrThrserProlys-----ValGlnArgAlaAsp 70  
 DB 203 -----AGAAAGCGGGTCCAAAGATTTGGAAACCATTAATAGAGTTGAA 247  
 QY 71 -----SerAspValIlePheSer-----76  
 DB 248 GTTGTGACACCTGAGAAACATTTGGGTGATGTATGTTGAATTTCTAGAAGAGA 307  
 QY 77 -----AnserPheGlyGluArgAsnValValThrGluGlyAspLeuLysVal 94  
 DB 308 CAATTAATAGCTTTGGTGACAAA-----CCTGGTGGCTC---AAGGTG 349  
 QY 95 LeuAspGlyCysAlaProLeuThr-----ArgPheThrAsnLysLeuArg-----109  
 DB 350 GTGGATCACTTGTGCCACTGACAGAAATGTTCAATATGAGAGCTTAAGGAATG 409  
 QY 110 ThrPheGlyArg-----ThrPheThrGluAlaTyralaAspPheCysIleAlaTyrlaHis 128  
 DB 410 ACMAAAGGTGAGCATCTTACCTATGCACTAGCAATGCAAAATTCATGATGTACTTCAAC 469  
 QY 129 LysLeuProGlnLeuAsnAlaAlaGluLeuGlyIleProAla 143  
 DB 470 ATTCAAGAACCACTTCTGCTGCCAAGAAAGAGAGTATACGCT 514

RESULT 15  
 LOCUS BG523153 648 bp mRNA linear EST 16-NOV-2001  
 DEFINITION 29-48 Stevia field grown leaf cDNA Stevia rebaudiana cDNA 5', mRNA  
 sequence.  
 ACCESSION BG523153  
 VERSION BG523153.1 GI:16946562  
 KEYWORDS EST.  
 SOURCE Stevia rebaudiana  
 ORGANISM Stevia rebaudiana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; campanulids; Asterales; Asteraceae; Asteroidae;

REFERENCE  
1 (bases 1 to 648)  
Eupatoriaceae; Stevia.  
AUTHORS  
Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.  
TITLE  
Leaf ESTs from Stevia rebaudiana: A resource for gene discovery in diterpene synthesis  
JOURNAL  
Unpublished (2001)  
COMMENT  
Contact: Jim Brandle  
Genomics and Biotechnology  
Agriculture and Agri-Food Canada - SCFRC  
1391 Sandford St., London, Ontario, CANADA, N5V 4T3  
Tel: 519 457 1470  
Fax: 519 457 3997  
Email: brandleje@em.agr.ca  
Seq primer: T3 promoter primer.  
Location/Qualifiers

## FEATURES

1..648  
/organism="Stevia rebaudiana"  
/mol\_type="mRNA"  
/strain="751/1501"  
/cultivar="Landrace"  
/db\_xref="taxon:55670"  
/tissue\_type="leaf"  
/dev\_stage="field grown, mid-size"  
/lab\_host="E. coli strain XL0LR"  
/clone\_lib="Stevia field grown leaf cDNA"  
/note="Vector: PBK-CMV; Site\_1: EcoRI; Site\_2: XhoI. This cDNA library was constructed from polyA+ enriched mRNA from field grown leaves. Mid-size actively growing leaves were collected and pooled from several plants and frozen immediately after harvesting in liquid nitrogen. The cDNA was prepared using an XhoI-poly(dT) linker-primer. An EcoRI adapter was ligated to the blunt end cDNA and the products were digested with EcoRI and XhoI enabling directional cloning into the lambda ZAP Express vector. The library was amplified using the host strain XL1-Blue MRF. Mass excision of the library was performed to obtain PBK-CMV phagemid clones in the host strain XL0LR. Single pass DNA sequencing was performed using the T3 promoter primer: 5' ATTACCTCCTCAAGGA 3'. This library was constructed by Alex Richman."

## ORIGIN

## Alignment Scores:

Pred. No.: 4.2 Length: 648  
Score: 84.50 Matches: 48  
Percent Similarity: 40.57% Conservative: 23  
Best Local Similarity: 27.43% Mismatches: 47  
Query Match: 8.53% Indels: 57  
DB: 12 Gaps: 12

US-09-613-486-15 (1-198) x BG523153 (1-648)

QY 2 GluLeuWetSerAspSerAsnLeuSerAsnLeu-----ValIle 14  
DB 80 GAGTCAATGCTTAATGAGTACTCGCGGCTTCCCTGTTGACCTTCGCGGTGTTG 139  
QY 15 ThrAspAlaSerSerLeuAsnGlyValAspLysLysLeuLeuSerAlaGluValGluLys 34  
DB 140 ACCGTT---GGCTCTACCATGATGATTTGACTCGAGTACTCGCGTCCAG----- 167  
QY 35 MetLeuValGlnLysGlyValAspAsnGluGlyLleGluValValPheGlyLeuLeuLeu 54  
DB 188 ---CTGGCGGCTCGAGGAGCTTCCGTCGAGAGATT----- 220  
QY 55 TyrAlaLeuAlaAlaArgThrTherProLys-----ValGlnArgAlaAsp 70  
DB 221 -----AGAAAGCGGGTCCCAAGATTTTGGAAACCATTAAGAGATTGA 265  
QY 71 -----SerAspValIlePheSer----- 76  
DB 266 GTTGTGACACCTGAGAACATTGGGTGATGTTGATTGATTGAATTTGAGAGAGGA 325  
QY 77 -----AsnSerPheGlyGluArgAsnValValValThrGluGlyAspLeuLysLysVal 94

DB 326 CAAATTAAATAGCTTTGGTGAACAA-----CTGGTGGCTCC---AAGGTG 367  
QY 95 LeuAspGlyCysAlaProLeuThr-----ArgPheThrAsnLysLeuArg----- 109  
DB 368 GTGGATTCACCTGTGGCCACTAGAGAAATGTTCAATATGTGAGTAAAGAGATG 427  
QY 110 ThrPheGlyArg---ThrPheThrGluAlaTyrValAspPheCysIleAlaTyrLysHis 128  
DB 428 ACAAAAGGTGAGCATCTTACACTATGCACTAGCAAAATTCATGTGATGCTCAACAC 487  
QY 129 LysLeuProGlnLeuAsnAlaAlaGluLeuGlyLleProAla 143  
DB 488 ATTGAGAACCACTTCTGCTGCCAAAGAAAGACAGTTACAGCT 532

Search completed: April 7, 2004, 03:41:23  
Job time : 2366 secs